



Filtrage linéaire par morceaux d'un système en temps discret avec petit bruit d'observation

Paula Milheiro de Oliveira, Marie-Christine Roubaud

► To cite this version:

Paula Milheiro de Oliveira, Marie-Christine Roubaud. Filtrage linéaire par morceaux d'un système en temps discret avec petit bruit d'observation. [Rapport de recherche] RR-1451, INRIA. 1991. inria-00077105

HAL Id: inria-00077105

<https://hal.inria.fr/inria-00077105>

Submitted on 29 May 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



UNITÉ DE RECHERCHE
INRIA-SOPHIA ANTIPOLIS

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
B.P.105
78153 Le Chesnay Cedex
France
Tél.: (1) 39 63 55 11

Rapports de Recherche

N° 1451

Programme 5
Traitement du Signal,
Automatique et Productique

FILTRAGE LINEAIRE PAR MORCEAUX D'UN SYSTEME EN TEMPS DISCRET AVEC PETIT BRUIT D'OBSERVATION

Paula MILHEIRO de OLIVEIRA
Marie-Christine ROUBAUD

Juin 1991



FILTRAGE LINEAIRE PAR MORCEAUX D'UN SYSTEME EN
TEMPS DISCRET AVEC PETIT BRUIT D'OBSERVATION (*)

DISCRETE TIME PIECEWISE LINEAR
FILTERING WITH SMALL OBSERVATION NOISE (*)

Paula Milheiro de Oliveira (**)

&

Marie Christine Roubaud

INRIA Sophia-Antipolis

BP 109

F 06561 VALBONNE Cedex

(*) Partiellement soutenu par USACCE dans le cadre du contrat DAJA 45-90-C-0008.

(**) et Faculdade de Engenharia da U. do Porto, Rua dos Bragas, 4099 PORTO Codex
(Portugal).

Résumé

Nous nous intéressons à un problème de filtrage linéaire par morceaux en temps discret avec petit bruit d'observation. Nous présentons et comparons plusieurs tests permettant de déterminer les intervalles de linéarité de la fonction d'observation, notamment dans le cas où elle est symétrique. Sur chacun de ces intervalles, nous approchons le filtre optimal par le filtre de Kalman-Bucy correspondant. Comme dans [FJSZ], nous approchons des processus discrets par des diffusions pour estimer les probabilités d'erreur et les temps moyens pour prendre une décision.

Abstract

We are interested in a piecewise linear discrete time filtering problem with small observation noise. We present and compare different tests which enable us to compute the intervals of linearity of the observation function, specially when it is symmetric. Over such an interval, we can then approximate the optimal filter by the corresponding Kalman-Bucy filter. As in [FJSZ], we approximate discrete time processes by diffusions, in order to estimate the probabilities of error and the expected times for taking a decision.

Table des matières

1	Introduction	1
2	Deux filtres de Kalman–Bucy en parallèle	5
3	Tests de détection des passages de $\{x_k\}$ par zéro	8
3.1	Test sur les observations y_k	8
3.2	Test sur la sortie d’un des filtres de Kalman–Bucy	10
4	Sous l’hypothèse (HD1): décider du signe de x_k	17
4.1	Test de la variation quadratique	17
4.2	Test du rapport de vraisemblance sur les sorties des filtres de Kalman–Bucy	20
5	Sous l’hypothèse (HD2): décider du signe de x_k	24
5.1	Test du rapport de vraisemblance sur les accroissements des observations	24
5.1.1	Le cas $\varepsilon = 0$	24
5.1.2	Le cas $\varepsilon > 0$	32
5.2	Test du rapport de vraisemblance sur les sorties des filtres de Kalman–Bucy	33
6	Application	37
6.1	Les critères de comparaison	37
6.2	Exemples	37
6.2.1	Exemples sous l’hypothèse (HD2)	38
6.2.2	Exemple sous l’hypothèse (HD1)	41
6.2.3	Un exemple sous l’hypothèse (HD2) ne vérifiant pas (H4)	42
7	Conclusion	43
A	Résultats numériques et représentations graphiques des exemples présentés dans la section 6	46
A.1	Résultats numériques	46
A.2	Représentations graphiques	49

1 Introduction

On s'intéresse au problème de filtrage non linéaire suivant:

$$\begin{cases} dX_t = b(X_t)dt + \sigma(X_t)dU_t \\ dY_t = h(X_t)dt + \varepsilon dV_t, \quad Y_0 = 0, \end{cases} \quad (1.1)$$

où $\{X_t, t \geq 0\}$ est le processus à valeurs dans \mathbb{R} non observé à estimer à l'instant τ au vu des observations jusqu'à l'instant τ du processus unidimensionnel $\{Y_t, t \geq 0\}$, ε un paramètre "petit" et $\{U_t, t \geq 0\}$ et $\{V_t, t \geq 0\}$ sont des processus de Wiener standards indépendants, à valeurs dans \mathbb{R} .

Il est bien connu que le problème est de dimension infinie, au sens où, pour le résoudre, on a à déterminer la solution d'une équation aux dérivées partielles, par exemple l'équation de Zakai. Si la fonction h est monotone, sous certaines hypothèses générales de "régularité", le filtre de Kalman étendu, entre autres, est une "bonne" approximation du filtre optimal (cf. [Picard], [KBS], [Bensoussan], [Ji] et [Milheiro]). Le problème avec h non monotone, sous une certaine "hypothèse de détectabilité" a été traité dans [Fleming-Pardoux].

Dans le cadre du filtrage linéaire par morceaux, on suppose que

- $b(x) = B_-x\mathbf{1}_{\{x < 0\}} + B_+x\mathbf{1}_{\{x \geq 0\}}$
- $\sigma(x) = \sigma_- \mathbf{1}_{\{x < 0\}} + \sigma_+ \mathbf{1}_{\{x \geq 0\}}$
- $h(x) = H_-x\mathbf{1}_{\{x < 0\}} + H_+x\mathbf{1}_{\{x \geq 0\}}$.

Si $H_+H_- > 0$, i.e. la fonction h est monotone, pour $\varepsilon = 0$, X_t peut être parfaitement connu et pour ε "petit", le problème n'offre pas de grandes difficultés.

Dans le cas $H_+H_- < 0$, i.e. h non monotone, bien que $h(X_t)$ puisse être estimé de façon précise, il n'est pas immédiat qu'il en soit de même pour X_t . Le filtre de Kalman étendu est en général inefficace, le problème étant de déterminer le signe de $\{X_t, t \geq 0\}$. La détermination du signe n'étant possible que dans le cas où la variance conditionnelle est petite ⁽¹⁾, on introduit l'hypothèse de "détectabilité" suivante notée (HD) :

$$(HD) \quad \begin{cases} H_-^2\sigma_-^2 \neq H_+^2\sigma_+^2 & (HD1) \\ \text{ou} \\ H_-^2\sigma_-^2 = H_+^2\sigma_+^2 \text{ et } B_+ \neq B_- & (HD2). \end{cases}$$

¹Contre-exemple: $b = 0, h(x) = |x|, \sigma = 1$.

Sous (HD1), dans [FJP], un filtre approché a été proposé et dans [Roubaud], il est montré que sous (HD2) un filtre du même type peut être utilisé comme approximation du filtre optimal.

L'idée est de construire un filtre approché à partir de deux filtres de Kalman-Bucy, associés respectivement aux deux problèmes de filtrage linéaire suivants:

$$\begin{cases} dX_t = B_+ X_t dt + \sigma_+ dU_t \\ dY_t = H_+ X_t dt + \varepsilon dU_t \end{cases} \quad (1.2)$$

et

$$\begin{cases} dX_t = B_- X_t dt + \sigma_- dU_t \\ dY_t = H_- X_t dt + \varepsilon dU_t. \end{cases} \quad (1.2)$$

On calcule ces deux filtres "en parallèle" et on utilise deux tests:

- Un premier test permet de détecter des intervalles de temps durant lesquels la trajectoire de $\{X_t\}_t$ ne passe pas par 0.
- Un second test permet de décider si $X_t < 0$ ou $X_t > 0$ sur ces intervalles, sous l'hypothèse (HD1) ou (HD2).

Sur chaque intervalle de monotonie, on approche alors le filtre optimal par le filtre de Kalman-Bucy correspondant.

Nous considérons l'intervalle de temps fini $[0, T]$.

Dans ce rapport, nous nous intéressons à la résolution numérique du problème linéaire par morceaux dans la situation où la fonction d'observation h est non monotone. Nous commençons par discrétiser le système continu (1.1) par un schéma classique de discrétisation en temps avec pas $\Delta t = \varepsilon$. Le processus $\{X_{k\Delta t}\}_k$ est approximé par $\{x_k\}_k$, $\{U_{(k+1)\Delta t} - U_{k\Delta t}\}_k$ par $\{\sqrt{\Delta t} u_k\}_k$, $\{V_{(k+1)\Delta t} - V_{k\Delta t}\}_k$ par $\{\sqrt{\Delta t} v_k\}_k$, $\{(\Delta t)^{-1} (Y_{(k+1)\Delta t} - Y_{k\Delta t})\}_k$ par $\{y_k\}_k$.

Nous obtenons le modèle discret suivant:

$$\begin{cases} x_{k+1} = x_k + \varepsilon b(x_k) + \sqrt{\varepsilon} \sigma(x_k) u_k \\ y_k = h(x_k) + \sqrt{\varepsilon} v_k, \end{cases} \quad (1.2)$$

avec les fonctions b , σ et h définies comme précédemment. On émet les hypothèses suivantes:

- (H1) $\{u_k\}_k$ et $\{v_k\}_k$ sont des bruits blancs gaussiens standards et indépendants,
- (H2) x_0 est une variable aléatoire réelle telle que $\mathbf{E}(\exp \{c_0 x_0^2\}) < +\infty$, pour un certain $c_0 > 0$,

(H3) $H_- H_+ < 0$ et $\sigma_- \sigma_+ \neq 0$.

Sans réduire la généralité du problème, on suppose que $h(x) \geq 0 \forall x \in \mathbb{R}$, i.e. on suppose que

(H3') $H_- < 0 : H_+ > 0$ et $\sigma_- \sigma_+ \neq 0$.

Nous adaptons l'étude faite dans le cas continu au cas discret. Notre but est de mettre en œuvre les tests permettant de séparer les intervalles de positivité et de négativité de $\{x_k\}_k$ et de comparer leur performance. Nous nous intéressons principalement à la situation où l'hypothèse (HD2) est vérifiée, le problème sous l'hypothèse (HD1) ayant été traité dans [FJSZ].

L'étude de l'évolution de la loi conditionnelle, obtenue de façon approchée par résolution de l'équation de Zakai discretisée (cf. [Le Gland]), nous aidera à interpréter les résultats obtenus.

Ce rapport est organisé comme il suit:

Dans le paragraphe 2, nous présenterons les deux filtres de Kalman-Bucy (2.5) et (2.6) utilisés et nous formulerons des remarques préliminaires. Dans le paragraphe 3 nous nous intéresserons à la détection des passages du signal par 0. En analogie avec le cas continu deux tests seront proposés, l'un basé sur les accroissements des observations (cf. [FJP]) et l'autre sur la sortie d'un des filtres de Kalman-Bucy (cf. [Roubaud]). Sachant que sur un intervalle de temps $[a, b]$ le signal ne passe pas par 0, avec une probabilité donnée, nous considérons un autre type de tests pour décider du signe de $\{x_k\}_k$ sur $[a, b]$. Sous l'hypothèse (HD1), des études ont été faites par [FJSZ]. Ces auteurs ont mis en œuvre deux tests différents: l'un basé sur la variation quadratique, l'autre du type rapport de vraisemblance basé sur les sorties des filtres de Kalman-Bucy. Nous résumerons brièvement ces procédures dans le paragraphe 4. Sous l'hypothèse (HD2), nous présenterons dans le paragraphe 5 un test de rapport de vraisemblance basé sur les accroissements des observations et adapterons le test sur les sorties des filtres de Kalman-Bucy du paragraphe précédent au cas traité.

La mise en œuvre des tests de détection et de décision du signe ainsi décrits nécessite la détermination de formules explicites pour les bornes. Bien qu'une démonstration rigoureuse n'ait pu être faite, les formules proposées sont justifiées de manière heuristique. Dans le cas où $B_- \neq B_+$, nous émettrons l'hypothèse supplémentaire:

(H4): $B_+ < 0$ et $B_- < 0$.

Cette hypothèse n'est pas trop restrictive pour notre propos. En effet, suivant le signe des coefficients de dérive, nous avons deux comportements distincts

du processus $\{x_k\}$. Si B_+ et B_- sont négatifs, il devient stationnaire dès que l'intervalle de monotonie est assez grand. Sinon, il peut "fuir" vers l'infini. En fait, nous étudions la situation la plus délicate au sens où la probabilité pour que le processus $\{x_k\}$ passe par zéro est importante. Cependant, si les valeurs $|B_-|$ et $|B_+|$ sont grandes, $\{x_k\}$ aura tendance à changer rapidement de signe et les tests n'auront probablement pas suffisamment de temps pour décider.

Le paragraphe 6 est consacré à l'application numérique de ces méthodes. Des critères de comparaison entre les différents tests seront proposés et les résultats des applications à divers exemples seront présentés. D'autre part, la solution de l'équation de Zakai sera utilisée pour la justification du comportement des filtres approchés.

Pour conclure, on présente dans le paragraphe 7 quelques réflexions sur le comportement de ces filtres et on discute la performance des différents tests étudiés.

Notation 1.1 *Etant donné un processus $\{G_k\}$ on écrira*

$$G_k = \mathcal{O}(\varepsilon^q),$$

où $q \in \mathbb{R}_+$, pour signifier qu'il existe $c_1, c_2, c_3 > 0$ tels que

$$E[G_k^2] \leq c_1 \exp\{-c_2 k \varepsilon\} + c_3 \varepsilon^{2q}, \forall k \in \mathcal{N}, \varepsilon > 0.$$

Notation 1.2 *On notera c ou C des constantes indépendantes de ε , sans se soucier de leurs valeurs. Ces constantes pourront donc être différentes d'une ligne à l'autre.*

2 Deux filtres de Kalman–Bucy en parallèle

Soit \mathcal{Y}_0^k , la tribu des observations jusqu'à l'instant k , $\mathcal{Y}_0^k = \sigma\{y_0, y_1, \dots, y_k\}$. La solution du problème de filtrage associé au système (1.1) étant donnée par la loi conditionnelle de x_k sachant \mathcal{Y}_0^k , nous commentons brièvement son comportement (cf. [FJSZ]). Sous l'hypothèse de "déteçtabilité" (HD), la variance de cette loi est "petite", sa densité se concentrant autour de deux maxima locaux, l'un négatif et l'autre positif. Ce phénomène se justifie du fait qu'on a, soit $x_k - y_k/H_- = \sqrt{\varepsilon}/H_- v_k$, soit $x_k - y_k/H_+ = \sqrt{\varepsilon}/H_+ v_k$, donc la densité conditionnelle est "petite" excepté sur deux intervalles d'amplitude d'ordre $\mathcal{O}(\sqrt{\varepsilon})$, autour de y_k/H_- et de y_k/H_+ . En émettant l'hypothèse de déteçtabilité (HD) nous nous plaçons dans la situation où lorsque $\{x_k\}$ prend ses valeurs loin de 0 durant un certain intervalle de temps, un des pics tend à disparaître, la loi conditionnelle s'approchant d'une loi gaussienne. Il apparaît alors légitime d'approcher le filtre optimal par un filtre de Kalman–Bucy. Intuitivement on s'attend à ce que, sous l'hypothèse ($HD2$), le temps nécessaire pour faire "disparaître" l'un des deux pics soit plus long que sous l'hypothèse ($HD1$). Cette idée sera illustrée dans le paragraphe 6 (voir les figures 3 à 6 et 12 à 14).

Soient (\hat{x}_k^+, Q^+) et (\hat{x}_k^-, Q^-) les filtres de Kalman–Bucy associés respectivement aux systèmes linéaires

$$\begin{cases} x_{k+1} &= (1 + B_+ \varepsilon) x_k + \sqrt{\varepsilon} \sigma_+ u_k \\ y_k &= H_+ x_k + \sqrt{\varepsilon} v_k, y_0 = 0 \end{cases} \quad (2.5)$$

et

$$\begin{cases} x_{k+1} &= (1 + B_- \varepsilon) x_k + \sqrt{\varepsilon} \sigma_- u_k \\ y_k &= H_- x_k + \sqrt{\varepsilon} v_k, y_0 = 0, \end{cases} \quad (2.5)$$

avec conditions initiales gaussiennes

- $\hat{x}_0^+ = \hat{x}_0^- = E(x_0)$
- $Q_0^+ = Q^+, Q_0^- = Q^-$

où Q^+ et Q^- sont les variances stationnaires des lois conditionnelles.

Remarque 2.1 Les expressions de ces variances se calculent facilement. (cf. [Milheiro]).

Les filtres sont donnés par les équations suivantes:

$$\begin{cases} \hat{x}_{k+1}^+ &= (1 + \varepsilon B_+) \hat{x}_k^+ + \frac{1}{\varepsilon} H_+ Q^+ (y_{k+1} - H_+ (1 + \varepsilon B_+) \hat{x}_k^+) \\ Q^+ &= \varepsilon \frac{\rho_+ + [\rho_+^2 + 4\sigma_+^2 H_+^2]^{\frac{1}{2}}}{|H_+| [2 + \rho_+ + [\rho_+^2 + 4\sigma_+^2 H_+^2]^{\frac{1}{2}}]} \end{cases} \quad (2.5)$$

avec

$$\rho_+ = (2B_+ + B_+^2 \varepsilon) \varepsilon + \sigma_+^2 H_+^2.$$

$$\begin{cases} \hat{x}_{k+1}^- &= (1 + \varepsilon B_-) \hat{x}_k^- + \frac{1}{\varepsilon} H_- Q^- (y_{k+1} - H_- (1 + \varepsilon B_-) \hat{x}_k^-) \\ Q^- &= \varepsilon \frac{\rho_- + [\rho_-^2 + 4\sigma_-^2 H_-^2]^{\frac{1}{2}}}{|H_-| [2 + \rho_- + [\rho_-^2 + 4\sigma_-^2 H_-^2]^{\frac{1}{2}}]} \end{cases} \quad (2.6)$$

avec

$$\rho_- = (2B_- + B_-^2 \varepsilon) \varepsilon + \sigma_-^2 H_-^2.$$

Remarque 2.2 Le fait d'avoir considéré une condition initiale gaussienne avec pour variance, la variance conditionnelle stationnaire, n'est pas gênant pour la suite de notre propos. En effet, puisque les systèmes étudiés se caractérisent par une "mémoire courte", l'influence de la condition initiale tend à disparaître rapidement et la variance conditionnelle évolue rapidement vers sa valeur stationnaire.

Les gains stationnaires sont donc,

$$K^+ = \frac{1}{\varepsilon} H_+ Q^+ \quad \text{et} \quad K^- = \frac{1}{\varepsilon} H_- Q^-.$$

On introduit les processus dits "d'innovation"

$$\nu_{k+1}^+ = y_{k+1} - H_+ (1 + \varepsilon B_+) \hat{x}_k^+$$

et

$$\nu_{k+1}^- = y_{k+1} - H_- (1 + \varepsilon B_-) \hat{x}_k^-.$$

Les processus ainsi définis sont considérés approximativement comme des "bruits blancs" au sens où leurs variances sont d'ordre $\mathcal{O}(\varepsilon)$, tandis que les corrélations $E(\nu_{k+1}^+ \nu_k^+)$ et $E(\nu_{k+1}^- \nu_k^-)$ sont d'ordre $\mathcal{O}(\varepsilon^2)$.

On notera $\varepsilon\vartheta^+$ et $\varepsilon\vartheta^-$ leurs variances respectives,

$$\begin{aligned}\varepsilon\vartheta^+ &= \varepsilon(1 + \sigma_+^2 H_+^2) + H_+^2(1 + \varepsilon B_+)^2 Q^+, \\ \varepsilon\vartheta^- &= \varepsilon(1 + \sigma_-^2 H_-^2) + H_-^2(1 + \varepsilon B_-)^2 Q^-.\end{aligned}$$

Dans la suite, on étudiera la procédure de test permettant de séparer les intervalles de monotonie de la fonction h et de décider du signe de x_k . Sur chacun de ces intervalles, il sera alors possible d'approcher le filtre optimal par le filtre de Kalman–Bucy correspondant.

3 Tests de détection des passages de $\{x_k\}$ par zéro

Nous devons premièrement déterminer sur l'intervalle de temps $[0, T]$, les intervalles sur lesquels $\{x_k\}$ ne passe pas par zéro, avec une probabilité donnée, proche de 1. Nous présentons deux tests de détection, le premier s'appliquant aux observations y_k , le second à la sortie de l'un des filtres de Kalman-Bucy.

Soit l'intervalle de temps $[a, b]$, avec $0 < a < b < T$. On pose $m = \lfloor (b - a)/\varepsilon \rfloor$ et $i_0 = \lfloor a/\varepsilon \rfloor$. On considère les deux événements suivants:

$$\begin{aligned} A_- &= \{x_k < 0; k = i_0, i_0 + 1, \dots, i_0 + m\}, \\ A_+ &= \{x_k > 0; k = i_0, i_0 + 1, \dots, i_0 + m\}. \end{aligned}$$

On remarque que

$$(A_- \cup A_+)^c = \{x_k x_{k+1} < 0 \text{ pour un certain } k, i_0 \leq k \leq i_0 + m\}.$$

3.1 Test sur les observations y_k

Etant donné que " $h(x) = 0$ ssi $x = 0$ ", $h(x_k)$ "petit" implique par continuité de l'application h , x_k proche de zéro. Cependant $h(x_k)$ n'est pas observé, mais pour ε "petit", il est approché par y_k .

Soit $c_{obs} > 0$ une constante à déterminer.

On définit l'événement test suivant:

$$\mathcal{C} = \{|y_k| \geq c_{obs}\}.$$

On montre de façon similaire à la version en temps continu (cf. [FJP][proposition 3.1]) la proposition suivante (cf. [FJSZ][proposition 2.1]):

Théorème 3.1 *Soit $c_{obs} > 0$ et $\varepsilon_0 > 0$ donnés, il existe $\beta > 0$ tel que $\forall \varepsilon \in (0, \varepsilon_0]$ on a*

$$P(\{(A_- \cup A_+)^c | \mathcal{C}\}) \leq e^{-\beta/\varepsilon}.$$

Détermination de la constante c_{obs} :

Nous utilisons, pour la détermination de la constante c_{obs} , un raisonnement similaire à celui proposé dans [FJSZ], sauf que dans notre cas le coefficient de dérive n'est plus supposé constant mais constant par morceaux. Intuitivement le choix de c_{obs} doit être un compromis entre "conserver des intervalles de monotonie suffisamment grands" et "réduire la probabilité d'erreur du test".

On cherche à ce que la probabilité

$$P(\{x_k x_{k+1} < 0\} \mid \{|y_k| \geq c_{obs}, |y_{k+1}| \geq c_{obs}\})$$

soit faible. D'après l'hypothèse (H3) on obtient:

$$P(\{y_k < -c_{obs}\}) \leq P(\{v_k < -c_{obs}/\sqrt{\varepsilon}\}) \leq \frac{1}{2} e^{-c_{obs}^2/(2\varepsilon)}.$$

Nous nous permettons donc de considérer uniquement le conditionnement par rapport à l'événement $\{y_k \geq c_{obs}, y_{k+1} \geq c_{obs}\}$. De plus, on a:

$$\{x_k x_{k+1} < 0\} \cap \{y_k \geq c_{obs}, y_{k+1} \geq c_{obs}\} = D_{c_{obs}}^1 \cup D_{c_{obs}}^2,$$

avec les définitions suivantes pour les évènements D_c^1, D_c^2 :

$$\begin{aligned} D_c^1 &= \{x_k < 0, x_{k+1} > 0\} \cap \{y_k \geq c, y_{k+1} \geq c\}, \\ D_c^2 &= \{x_k > 0, x_{k+1} < 0\} \cap \{y_k \geq c, y_{k+1} \geq c\}, \end{aligned}$$

pour tout $c > 0$.

On considère séparément les évènements $D_{c_1}^1$ et $D_{c_2}^2$.

• Sur $D_{c_1}^1$, on a les récurrences suivantes:

$$\begin{aligned} x_{k+1} &= (1 + \varepsilon B_-)x_k + \sigma_- \sqrt{\varepsilon} u_k, \\ y_k &= H_- x_k + \sqrt{\varepsilon} v_k, \quad y_{k+1} = H_+ x_{k+1} + \sqrt{\varepsilon} v_{k+1}. \end{aligned}$$

On déduit l'égalité:

$$\frac{1}{\sqrt{\varepsilon}}(H_- y_{k+1} - H_+(1 + \varepsilon B_-)y_k) = H_+ H_- \sigma_- u_k + H_- v_{k+1} - H_+(1 + \varepsilon B_-)v_k.$$

Or, d'après (H3') on a $h(x_k) \geq 0$ et, pour ε suffisamment petit, on obtient la majoration

$$\frac{1}{\sqrt{\varepsilon}} |H_- - H_+(1 + \varepsilon B_-)| c_1 \leq |Z_k|,$$

où

$$Z_k = H_+ H_- \sigma_- u_k + H_- v_{k+1} - H_+(1 + \varepsilon B_-)v_k$$

suit une loi gaussienne $\mathcal{N}(0, H_+^2 H_-^2 \sigma_-^2 + H_-^2 + H_+^2 (1 + \varepsilon B_-)^2)$.

Soit $\{Z_k^0\}$ le processus normalisé associé à $\{Z_k\}$. Pour un risque α donné ⁽²⁾, on détermine $\lambda > 0$ tel que

$$P(|Z_k^0| \geq \lambda) \leq \alpha,$$

²Généralement on prend $\alpha = 0.05$.

par inversion de la fonction de répartition de la loi normale centrée réduite. On obtient ainsi l'expression de la constante c_1 :

$$c_1 = \lambda \sqrt{\varepsilon} \frac{\sqrt{H_+^2 H_-^2 \sigma_-^2 + H_-^2 + H_+^2 (1 + \varepsilon B_-)^2}}{|H_- - H_+ (1 + \varepsilon B_-)|}.$$

• Sur $D_{c_2}^2$, un raisonnement similaire au précédent nous amène à considérer une constante c_2 donnée par:

$$c_2 = \lambda \sqrt{\varepsilon} \frac{\sqrt{H_+^2 H_-^2 \sigma_+^2 + H_+^2 + H_-^2 (1 + \varepsilon B_+)^2}}{|H_+ - H_- (1 + \varepsilon B_+)|}.$$

Dans les applications numériques, pour une probabilité d'erreur α donnée, nous prendrons $c_{obs} = \max(c_1, c_2)$.

Remarque 3.2 Dans les calculs de la constante c_{obs} , la prise en compte de la dérive introduit uniquement des termes d'ordre $\mathcal{O}(\varepsilon^{3/2})$ que nous négligerons devant les termes d'ordre $\mathcal{O}(\sqrt{\varepsilon})$.

3.2 Test sur la sortie d'un des filtres de Kalman–Bucy

On se propose de décrire un test capable de détecter les changements de signe éventuels du processus $\{x_k\}$ à l'aide de la sortie obtenue par un des deux filtres de Kalman–Bucy, par exemple $\{\hat{x}_k^+\}$. La construction d'un tel test se justifie par le fait que “ $h(x) = 0$ ssi $x = 0$ ” et par la proposition suivante:

Proposition 3.3 Soit $0 < a < b$. Pour tout $\theta > 0$, il existe $\beta > 0$, $\varepsilon_0 > 0$ tels que, pour $\varepsilon \in (0, \varepsilon_0]$, si on définit $i_0 = \lfloor a/\varepsilon \rfloor$ et $m = \lfloor (b-a)/\varepsilon \rfloor$, on a:

$$P(\{\max_{k=i_0, \dots, i_0+m} |h(x_k) - H_+ \hat{x}_k^+| > \theta\}) \leq e^{-\beta/\sqrt{\varepsilon}}.$$

En vu de la démonstration de cette proposition, on établit d'abord le lemme suivant:

Lemme 3.4 Pour tout $T > 0$, il existe $\varepsilon_0 > 0$ et $c, C > 0$ tels que, pour tout $\varepsilon \in (0, \varepsilon_0]$, $m = \lfloor T/\varepsilon \rfloor$,

$$\mathbf{E}(\exp \max_{k=0, \dots, m} c x_k^2) \leq C.$$

Preuve

La preuve de ce lemme est présentée en 3 étapes (cf. [Roubaud][lemme I.2.7] pour la preuve analogue en temps continu).

Etape 1:

On démontre le résultat suivant:

Lemme 3.5 *On définit le processus $\{z_k\}$ par*

$$z_{k+1} = (1 + \varepsilon B)z_k + \sqrt{\varepsilon} \sigma u_k, \quad (3.11)$$

où z_0 est une variable aléatoire indépendante du processus $\{u_k\}$ et il existe des constantes $c_0, C_0 > 0$ telles que

$$\mathbf{E}(\exp c_0 z_0^2) < C_0. \quad (3.11)$$

Alors, pour tout $T > 0$, il existe $\varepsilon_0 > 0$ et $c, C > 0$ tels que, pour tout $\varepsilon \in (0, \varepsilon_0]$, $m = \lceil T/\varepsilon \rceil$,

$$\mathbf{E}(\exp \max_{k=0, \dots, m} c z_k^2) \leq C.$$

Preuve

Nous réécrivons l'expression de z_{k+1} en mettant en évidence une partie dépendante de z_0 et une partie martingale:

$$z_{k+1} = (1 + \varepsilon B)^k [(1 + \varepsilon B)z_0 + M_k],$$

avec

$$M_k = \sigma \sqrt{\varepsilon} \sum_{j=0}^k (1 + \varepsilon B)^{-j} u_j.$$

$$z_{k+1}^2 \leq (1 + \varepsilon B)^{2k} [2(1 + \varepsilon B)^2 z_0^2 + 2M_k^2] \leq 2e^{2TB} [(1 + \varepsilon B)^2 z_0^2 + M_k^2]$$

D'après (3.11) et l'inégalité de Cauchy-Schwartz, il est suffisant de montrer qu'il existe $\varepsilon_0, c, C > 0$ tels que $\forall \varepsilon \in (0, \varepsilon_0]$,

$$\mathbf{E}(\exp \max_{k=0, \dots, m-1} c M_k^2) \leq C.$$

Or $\{M_k\}$ est un processus gaussien centré de variance $\mathbf{E}[M_k^2]$,

$$\begin{aligned} \mathbf{E}[M_k^2] &= (k+1) \varepsilon \sigma^2 & \text{si } B = 0, \\ \mathbf{E}[M_k^2] &= \sigma^2 \frac{(1 + \varepsilon B)^2 - (1 + \varepsilon B)^{-2k}}{B(2 + \varepsilon B)} & \text{si } B \neq 0. \end{aligned}$$

$$\text{Si } B > 0, \quad \mathbf{E}[M_k^2] \leq \sigma^2 \frac{(1 + \varepsilon B)^2}{B(2 + \varepsilon B)}$$

$$\text{Si } B < 0, \quad \mathbf{E}[M_k^2] \leq \sigma^2 \frac{-(1 + \varepsilon B)^2 + \exp(2k\varepsilon B + k\varepsilon^2 B^2)}{-B(2 + \varepsilon B)}$$

Il existe donc ε_0 et une constante $C = C(\varepsilon_0, T) > 0$ tels que $\forall \varepsilon \in (0, \varepsilon_0]$,

$$\max_{k=0, \dots, m-1} \mathbf{E}[M_k^2] \leq C.$$

On en déduit l'existence de $c_1 > 0$ tel que

$$\limsup_{\varepsilon \rightarrow 0} \mathbf{E}(\exp c_1 M_{m-1}^2) \leq C.$$

Vu que $\{\exp c M_k^2\}$ est une sous-martingale, on a:

$$\mathbf{E}(\max_{k=0, \dots, m-1} \exp c M_k^2) \leq 4 \mathbf{E}(\exp c_1 M_{m-1}^2) \leq C,$$

pour $c \leq c_1/2$, d'où le lemme. ■

Etape 2:

On introduit les notations:

$$B = \sup(|B_+|, |B_-|)$$

$$\sigma = \sup(|\sigma_+|, |\sigma_-|)$$

On démontre que les moments d'ordre $2p$, $p \in \mathcal{N}$, du processus $\{x_k\}$ peuvent être majorés par les moments d'ordre $2p$ du processus $\{z_k\}$ défini par

$$\begin{cases} z_{k+1} = (1 + \varepsilon B)z_k + \sqrt{\varepsilon} \sigma u_k, \\ z_0 = x_0. \end{cases}$$

En effet, d'après l'équation d'état, on a

$$x_{k+1}^{2p} = \sum_{j=0}^{2p} C_j^{2p} [x_k + \varepsilon b(x_k)]^j [\sqrt{\varepsilon} \sigma(x_k) u_k]^{2p-j}.$$

Le fait que $\{u_k\}$ est un bruit blanc et que u_k est indépendant de x_k , entraîne l'égalité

$$\mathbf{E}[x_{k+1}^{2p}] = \sum_{j=0; j \text{ pair}}^{2p} C_j^{2p} \mathbf{E}[(x_k + \varepsilon b(x_k))^j (\sqrt{\varepsilon} \sigma(x_k))^{2p-j}] \mathbf{E}[u_k^{2p-j}].$$

On obtient la majoration suivante:

$$\mathbf{E}[x_{k+1}^{2p}] \leq \sum_{j=0; j \text{ pair}}^{2p} C_j^{2p} (\sqrt{\varepsilon} \sigma)^{2p-j} \frac{(2p-j)!}{2^{p-j/2} (p-j/2)!} (1 + \varepsilon B)^j \mathbf{E}[x_k^j].$$

D'autre part, on note que

$$\mathbf{E}[z_{k+1}^{2p}] = \sum_{j=0; j \text{ pair}}^{2p} C_j^{2p} (\sqrt{\varepsilon} \sigma)^{2p-j} \frac{(2p-j)!}{2^{p-j/2} (p-j/2)!} (1 + \varepsilon B)^j \mathbf{E}[z_k^j].$$

Vu que $z_0 = x_0$, on obtient

$$\mathbf{E}[x_k^{2p}] \leq \mathbf{E}[z_k^{2p}], \quad k = 0, 1, \dots, [T/\varepsilon].$$

Le lemme 3.5 entraine l'existence de $\varepsilon_0 > 0$ et $c, C > 0$ tels que $\forall \varepsilon \in (0, \varepsilon_0]$,

$$\max_{k=0, \dots, m} \mathbf{E}(\exp c x_k^2) \leq C. \quad (3.13)$$

Etape 3:

Nous utilisons la majoration (3.13) pour obtenir le lemme 3.4. D'après la formule de Taylor, on a:

$$\begin{aligned} \exp c x_{j+1}^2 &= \exp c x_j^2 + 2c x_j \exp(c x_j^2) (x_{j+1} - x_j) \\ &\quad + 2c \exp(c \theta_j^2) (1 + 2c \theta_j^2) (x_{j+1} - x_j)^2. \end{aligned}$$

On obtient l'expression

$$\begin{aligned} \exp c x_{k+1}^2 &= \exp c x_0^2 + 2c \sum_{j=0}^k x_{k-j} \exp(c x_{k-j}^2) (x_{k+1-j} - x_{k-j}) \\ &\quad + 2c \sum_{j=0}^k \exp(c \theta_{k-j}^2) (1 + 2c \theta_{k-j}^2) (x_{k+1-j} - x_{k-j})^2. \end{aligned}$$

Puisque $|\theta_k| \leq |x_k| + |x_{k+1}|$, on a la majoration

$$\begin{aligned} &\mathbf{E} \left[\max_{k=0, \dots, m-1} \exp c x_{k+1}^2 \right] \\ &\leq \mathbf{E}[\exp c x_0^2] + 2c \varepsilon B \sum_{j=0}^{m-1} \mathbf{E}[x_{k-j}^2 \exp(c x_{k-j}^2)] \\ &\quad + 2c \mathbf{E} \left[\max_{k=0, \dots, m-1} \sqrt{\varepsilon} \left(\sum_{j=0}^k \exp(c \theta_{k-j}^2) x_{k-j} \sigma(x_{k-j}) u_{k-j} \right) \right] \\ &\quad + 4c \varepsilon \sum_{j=0}^{m-1} \mathbf{E} \left[\exp(2c(x_{k-j}^2 + x_{k+1-j}^2)) (1 + 4c(x_{k+1-j}^2 + x_{k-j}^2)) \right. \\ &\quad \left. (\varepsilon B^2 x_{k-j}^2 + \sigma^2 u_{k-j}^2) \right]. \quad (3.13) \end{aligned}$$

En utilisant l'hypothèse (H2), l'inégalité de Cauchy-Schwartz et le résultat (3.13) nous remarquons que le lemme est démontrée s'il existe $\varepsilon_0 > 0$ et $c, C > 0$ tels que $\forall \varepsilon \in (0, \varepsilon_0]$,

$$\mathbb{E} \left[\max_{k=0, \dots, m-1} \sqrt{\varepsilon} \left(\sum_{j=0}^k \exp(c, x_{k-j}^2) x_{k-j} \sigma(x_{k-j}) u_{k-j} \right) \right] \leq C.$$

Le processus

$$M_k = \sqrt{\varepsilon} \sum_{j=0}^k \exp(c x_{k-j}^2) x_{k-j} \sigma(x_{k-j}) u_{k-j}$$

est une \mathcal{F}_k -martingale avec $\mathcal{F}_k = \sigma(x_0, u_0, \dots, u_k)$.

On utilise une propriété des martingales:

$$\mathbb{E} \left[\max_{k=0, \dots, m-1} M_k^2 \right] \leq 4 \mathbb{E}[M_{m-1}^2].$$

Vu que u_k et x_k sont indépendants et x_k est \mathcal{F}_{k-1} -mesurable, on a donc

$$\begin{aligned} \mathbb{E}[M_{m-1}^2] &\leq m \varepsilon \sigma \max_{j=0, \dots, m-1} \mathbb{E}[\exp(4 c x_{m-1-j}^2)]^{1/2} \mathbb{E}[x_{m-1-j}^4]^{1/2} \\ &\leq T c, \end{aligned}$$

d'après l'inégalité (3.13), d'où le lemme 3.4. ■

Preuve (de la proposition 3.3)

Nous utilisons la notion de sous-différentiel de h .

Puisque la fonction h est convexe, il existe ρ_k appartenant au sous-différentiel de h à un point θ_k tel que

$$h(x_{k+1}) = h(x_k) + \rho_k(x_{k+1} - x_k)$$

et

$$|\rho_k| \leq \sup(|H_+|, |H_-|) \triangleq H.$$

On définit

$$Z_k \triangleq h(x_k) - H_+ \hat{x}_k^+.$$

D'après l'équation du filtre \hat{x}_k^+ et l'équation d'état, on obtient l'expression récurrente:

$$Z_{k+1} = \delta Z_k + \varepsilon b(x_k) \rho_k (1 - H_+ K_+) + \sqrt{\varepsilon} \sigma(x_k) \rho_k (1 - H_+ K_+) u_k + \sqrt{\varepsilon} H_+ K_+ v_{k+1},$$

où

$$\delta \triangleq (1 + \varepsilon B_+) (1 - H_+ K_+).$$

On remarque que $0 < \delta < 1$. On utilise cette récurrence jusqu'au terme Z_{i_0} . On établit ainsi la majoration:

$$|Z_{k+1}| \leq \exp\{-c(k+1)\}|Z_{i_0}| + c\varepsilon \sum_{j=0}^{k-i_0} \delta^j |x_{k-j}| + c\sqrt{\varepsilon} \sum_{j=0}^{k-i_0} \delta^j |u_{k-j}| + c\sqrt{\varepsilon} \sum_{j=0}^{k-i_0} \delta^j |v_{k+1-j}|. \quad (3.14)$$

On traite les quatre termes séparément. L'inégalité de Bienaymé-Tchebichev et le fait que $\mathbf{E}[|Z_{i_0}|] < c$ avec $i_0 = \left\lceil \frac{a}{\varepsilon} \right\rceil$ implique

$$P\left(|Z_{i_0}| > \frac{\theta}{4} e^{ci_0}\right) \leq \frac{C}{\theta} e^{-ca/\varepsilon}.$$

D'autre part, en utilisant le lemme 3.4, on obtient la majoration

$$P\left(\varepsilon \sum_{j=0}^{k-i_0} \delta^j |x_{k-j}| > \frac{\theta}{4c}\right) \leq C e^{-\theta(1-\delta)/(4c\varepsilon)}$$

Le lemme 2.3. dans [Roubaud] entraîne la majoration

$$\begin{aligned} & P\left(\sqrt{\varepsilon} \sum_{j=0}^{k-i_0} \delta^j |u_{k-j}| > \frac{\theta}{4c}\right) \\ & \leq P\left(\max_{k=i_0, \dots, i_0+m-1} |u_k| > \frac{\theta(1-\delta)}{4c\sqrt{\varepsilon}}\right) \leq \frac{b-a}{\varepsilon} e^{-c\theta/\varepsilon}, \end{aligned}$$

donc il existe $\varepsilon_0 > 0$ tel que $\forall \varepsilon \in (0, \varepsilon_0]$,

$$P\left(\sqrt{\varepsilon} \sum_{j=0}^{k-i_0} \delta^j |u_{k-j}| > \frac{\theta}{4c}\right) \leq e^{-c/\varepsilon}.$$

De même pour le 4^{me} terme à droite dans l'expression (3.14), on a:

$$P\left(\sqrt{\varepsilon} \sum_{j=0}^{k-i_0} \delta^j |v_{k-j}| > \frac{\theta}{4c}\right) \leq e^{-c/\varepsilon}.$$

De l'expression (3.14), en utilisant les quatre majorations ci-dessus, on obtient la proposition 3.3. ■

Soit c_{FK} une constante strictement positive à déterminer.

On définit l'évènement test suivant:

$$\hat{C} = \{|\hat{x}_k^+| \geq c_{FK}; k = i_0, i_0 + 1, \dots, i_0 + m\}.$$

Comme conséquence de la proposition 3.3, on obtient le théorème suivant:

Théorème 3.6 Il existe $\beta > 0$, $\varepsilon_0 > 0$ tels que, pour $\varepsilon \in (0, \varepsilon_0]$, on a:

$$P((A_+ \cup A_-)^c | \hat{C}) \leq e^{-\beta/\sqrt{\varepsilon}}.$$

Preuve La preuve de ce théorème suit pas à pas celle du résultat équivalent en temps continu (cf. [Roubaud] [Théorème 3.4]).

■

Détermination de la constante c_{FK} :

On cherche une constante c_{FK} telle que la probabilité

$$P(\{x_k x_{k+1} \leq 0\} | \{|\hat{x}_k^+| \geq c_{FK}, |\hat{x}_{k+1}^+| \geq c_{FK}\})$$

soit faible.

On considère y_k comme une approximation de $h(x_k)$. Dans l'étude précédente, pour une probabilité d'erreur α donnée, on a déterminé une constante c_{obs} permettant de tester au vu de y_k , l'occurrence d'un changement du signe de x_k . Or, un changement de signe de x_k est directement lié au fait que $h(x_k)$ devient "petit". Si $H_+ \hat{x}_k^+$ était une meilleure approximation de $h(x_k)$ que y_k , tout au moins aux alentours de zéro, la probabilité d'erreur du test " $\hat{x}_k^+ \geq c_{obs}/H_+$ " serait minorée par α . Nous ne savons pas si cette condition est vérifiée, mais, sur l'évènement A_+ , le filtre \hat{x}_m^+ est une meilleure approximation de x_m que y_m/H_+ . Et, en tout cas, la proposition 3.3 nous assure que l'écart $|h(x_k) - H_+ \hat{x}_k^+|$ est petit avec une probabilité proche de 1. Il semble alors raisonnable de penser que le test " $\hat{x}_k^+ \geq c_{obs}/H_+$ " permettra de détecter convenablement les passages de x_k par 0. Nous proposons donc pour constante

$$c_{FK} = c_{obs} / |H_+|.$$

Le même type de raisonnement est valable quand le test est appliqué sur le filtre \hat{x}_k^- au lieu du filtre \hat{x}_k^+ . Nous prenons lors

$$c_{FK} = c_{obs} / |H_-|.$$

Sans perte de généralité on considèrera par la suite un seul intervalle $[a, b]$ représentant l'intervalle de temps durant lequel il n'y a pas de passage à 0, avec une probabilité d'erreur donnée, i_0 et m seront définis comme en début de cette section.

4 Sous l'hypothèse (HD1): décider du signe de x_k

Dans ce paragraphe, nous présentons deux tests permettant de choisir entre les événements A_+ et A_- , dans le cas où le problème de filtrage linéaire par morceaux vérifie l'hypothèse (HD1). Ces tests ont été étudiés dans [FJSZ]. Le problème analogue en temps continu a été traité dans [FJP]. Le premier test considéré, dit de la variation quadratique, n'est applicable que sous l'hypothèse (HD1). Le second test, du type rapport de vraisemblance sur les sorties du filtre de Kalman-Bucy, pourra être adapté au cas où l'hypothèse (HD2) est vérifiée (cf. paragraphe 5.2). Le premier test sera donc introduit de façon brève. Le deuxième test sera décrit de façon plus détaillée, puisque la même idée sera utilisée par la suite.

Dans [FJSZ], les auteurs ont mis en œuvre ces tests sous deux formes: test à taille d'échantillon fixée et test séquentiel. D'après leur étude, les tests séquentiels semblent plus intéressants que ceux à taille d'échantillon fixée, vis-à-vis du critère du "temps moyen pour prendre une décision". Nous ne présenterons donc que les tests séquentiels.

4.1 Test de la variation quadratique

On considère un test d'hypothèses basé sur la variation quadratique de la suite d'observations $\{y_k\}$ pour décider entre les deux alternatives " A_+ " et " A_- ". Cette décision est prise sur N observations ($0 < N \leq m$) dans un intervalle de monotonie obtenu par un des tests introduits dans le paragraphe 3. On considère N comme étant un temps d'arrêt.

Sans restreindre la généralité du problème, on suppose que $B_+ = B_- = B$.

Remarque 4.1 Le fait de considérer B_+ et B_- distincts a pour unique conséquence l'introduction de termes négligeables dans les expressions, la procédure restant essentiellement la même.

Notons:

$$\begin{aligned}\Delta_k &= y_{k+1} - (1 + B\varepsilon) y_k, \\ \Upsilon_+^2 &= H_+^2 \sigma_+^2 + 1 + (1 + B\varepsilon)^2, \\ \Upsilon_-^2 &= H_-^2 \sigma_-^2 + 1 + (1 + B\varepsilon)^2.\end{aligned}$$

Alors

- si $x_k \geq 0$ et $x_{k+1} \geq 0$, la variable aléatoire Δ_k suit une loi gaussienne de moyenne nulle et de variance $\Upsilon_+^2 \varepsilon$

- si $x_k < 0$ et $x_{k+1} < 0$, la variable aléatoire Δ_k suit une loi gaussienne de moyenne nulle et de variance $\Upsilon_-^2 \varepsilon$.

L'hypothèse de détectabilité (HD1) implique

$$\Upsilon_+^2 \neq \Upsilon_-^2.$$

Le problème de décider entre les alternatives " A_+ " et " A_- " conduit à un test d'hypothèses sur ces variances.

On définit le rapport logarithmique

$$Z_k = \ln \frac{\mathcal{N}(\Delta_k, 0, \Upsilon_+^2 \varepsilon)}{\mathcal{N}(\Delta_k, 0, \Upsilon_-^2 \varepsilon)},$$

où $\mathcal{N}(x, \mu, \gamma^2)$ représente la densité de la loi gaussienne de paramètres μ et γ^2 et la série

$$S_n = \sum_{k=i_0}^{i_0+n} Z_k,$$

où $0 \leq n \leq m$, i_0 et m étant définis dans le paragraphe 3.

On obtient les expressions suivantes:

$$Z_k = \log \frac{\Upsilon_-}{\Upsilon_+} + \frac{1}{2} \left(\frac{1}{\Upsilon_-^2} - \frac{1}{\Upsilon_+^2} \right) \frac{\Delta_k^2}{\varepsilon}$$

et

$$S_n = (n+1) \log \frac{\Upsilon_-}{\Upsilon_+} + \frac{1}{2} \left(\frac{1}{\Upsilon_-^2} - \frac{1}{\Upsilon_+^2} \right) \sum_{k=i_0}^{i_0+n} \frac{\Delta_k^2}{\varepsilon}.$$

Soit $l_1 > 0$ et $l_2 > 0$ fixés, on introduit le temps d'arrêt N^* :

$$N^* = \inf \{n > 0 : S_n < -l_1 \text{ ou } S_n > l_2\} \wedge m.$$

On décrit le test comme suit: si $S_{N^*} \geq l_2$ on accepte " A_+ " et si $S_{N^*} \leq -l_1$ on accepte " A_- "; sinon on dira que le test ne permet pas de décider.

Il reste à proposer une méthode de détermination des bornes l_1 et l_2 . On approche le processus discret $\{S_n\}$ par un certain processus de diffusion $\{\zeta_t\}$ dont on écrit les équations sous " A_+ " et sous " A_- ". On obtient ainsi, par la formule de Dynkin, des équations différentielles ordinaires permettant de calculer les temps moyens pour prendre une décision $E(T_+^*)$ et $E^-(T_-^*)$ et pour les probabilités d'erreur p_+ et p_- (cf. [Fleming-Rishel]).

On introduit les notations

$$\mu_+ = \log \frac{\Upsilon_-}{\Upsilon_+} + \frac{1}{2} \left(\frac{1}{\Upsilon_-^2} - \frac{1}{\Upsilon_+^2} \right) \Upsilon_+^2$$

$$\mu_- = \log \frac{\Upsilon_-}{\Upsilon_+} + \frac{1}{2} \left(\frac{1}{\Upsilon_-^2} - \frac{1}{\Upsilon_+^2} \right) \Upsilon_-^2$$

$$\gamma_+ = \frac{1}{2} \left(\frac{1}{\Upsilon_-^2} - \frac{1}{\Upsilon_+^2} \right) \sqrt{\Upsilon_+^4 + 2(1 + B\varepsilon)^2}$$

$$\gamma_- = \frac{1}{2} \left(\frac{1}{\Upsilon_-^2} - \frac{1}{\Upsilon_+^2} \right) \sqrt{\Upsilon_-^4 + 2(1 + B\varepsilon)^2}.$$

La résolution des équations nous donne comme solution:

$$p_+ = \frac{1 - e^{-\mu_+/\gamma_+^2 l_2}}{e^{\mu_+/\gamma_+^2 l_1} - e^{-\mu_+/\gamma_+^2 l_2}} \quad (4.19)$$

$$p_- = \frac{1 - e^{-\mu_-/\gamma_-^2 l_1}}{e^{-\mu_-/\gamma_-^2 l_2} - e^{\mu_-/\gamma_-^2 l_1}} \quad (4.19)$$

$$E(T_+^*) = \left(\frac{l_2}{\mu_+} - p_+ \frac{l_2 + l_1}{\mu_+} \right) \varepsilon$$

$$E(T_-^*) = \left(-\frac{l_1}{\mu_-} + p_- \frac{l_2 + l_1}{\mu_-} \right) \varepsilon,$$

où

- p_+ est "la probabilité de refuser " A_+ " à tort",
- p_- est "la probabilité de refuser " A_- " à tort",
- T_+^* et T_-^* sont les temps d'arrêt

$$T_+^* = \inf\{t : \zeta_t \geq \varepsilon l_2 \text{ ou } \zeta_t \leq -\varepsilon l_1\}, \text{ dans le cas positif,}$$

et

$$T_-^* = \inf\{t : \zeta_t \geq \varepsilon l_2 \text{ ou } \zeta_t \leq -\varepsilon l_1\}, \text{ dans le cas négatif.}$$

Pour des probabilités d'erreur p_+ et p_- données ⁽³⁾, les expressions (4.19) et (4.19) permettent de calculer les bornes l_1 et l_2 .

³Généralement on prend $p_+ = p_- = 0.05$.

4.2 Test du rapport de vraisemblance sur les sorties des filtres de Kalman-Bucy

On décrit un autre test du type rapport de vraisemblance pour décider entre les deux alternatives " A_+ " et " A_- ". L'idée de ce test est la suivante:

Soit i_1 un entier, $i_0 < i_1 < i_0 + m$. Intuitivement, on s'attend à ce que $\hat{h}_k = E(h(X_k)|\mathcal{Y}_0^k)$ soit très proche de $H_+ \hat{x}_k^+$ sur A_+ et de $H_- \hat{x}_k^-$ sur A_- , pour $i_1 \leq k \leq i_0 + m$. On a

$$\begin{aligned} \text{sur } A_+ : \quad y_k &= H_+(1 + \varepsilon B_+) \hat{x}_{k-1}^+ + \nu_k^+, \\ \text{sur } A_- : \quad y_k &= H_-(1 + \varepsilon B_-) \hat{x}_{k-1}^- + \nu_k^-, \end{aligned}$$

où $\{\nu_k^+\}$ et $\{\nu_k^-\}$ peuvent être considérés comme des bruits blancs gaussiens dits "processus d'innovation" (cf. section 2).

On pose

$$Z_k = H_+(1 + \varepsilon B_+) \hat{x}_k^+ - H_-(1 + \varepsilon B_-) \hat{x}_k^-$$

et on prend pour statistique de test le rapport logarithmique suivant:

$$\hat{L}_n = \frac{1}{\varepsilon} \left[\sum_{k=i_1}^{i_1+n} Z_k y_{k+1} - \frac{1}{2} \sum_{k=i_1}^{i_1+n} \left((H_+(1 + \varepsilon B_+) \hat{x}_k^+)^2 - (H_-(1 + \varepsilon B_-) \hat{x}_k^-)^2 \right) \right],$$

avec $i_1 \leq n \leq m$.

La statistique \hat{L}_n se réécrit sous la forme suivante:

$$\begin{aligned} \text{Sur } A_+ : \hat{L}_n &= +\frac{1}{2\varepsilon} \sum_{k=i_1}^{i_1+n} Z_k^2 + \frac{1}{\varepsilon} \sum_{k=i_1}^{i_1+n} Z_k \nu_{k+1}^+, \\ \text{Sur } A_- : \hat{L}_n &= -\frac{1}{2\varepsilon} \sum_{k=i_1}^{i_1+n} Z_k^2 + \frac{1}{\varepsilon} \sum_{k=i_1}^{i_1+n} Z_k \nu_{k+1}^-. \end{aligned}$$

On note \hat{R}_n le terme de précision du test:

$$\hat{R}_n = \frac{1}{\varepsilon} \sum_{k=i_1}^{i_1+n} Z_k^2.$$

Pour R_n "suffisamment grand", le signe de \hat{L}_n permet de choisir entre les alternatives " A_- " et " A_+ ".

Soit $l_1 > 0$ et $l_2 > 0$ des bornes fixées, on introduit le temps d'arrêt N^* :

$$N^* = \inf \{n \geq i_1 : \hat{L}_n \geq l_2 \text{ ou } \hat{L}_n \leq -l_1\} \wedge m.$$

On applique la règle de décision qui suit.

Règle de décision:

- Si $\hat{L}_{N^*} \geq l_2$, on décide " A_+ ".
- Si $\hat{L}_{N^*} \leq -l_1$, on décide " A_- ".
- Sinon, on ne peut pas décider.

Il nous reste à calculer les bornes l_1 et l_2 pour une probabilité d'erreur donnée α et à estimer les temps moyens pour atteindre une décision. Le raisonnement présenté par la suite constitue une justification purement heuristique du choix des constantes l_1 et l_2 . Il n'a pas la prétention d'être une preuve rigoureuse.

Si on se place dans le cadre de l'hypothèse " A_+ ", d'après les équations des filtres de Kalman-Bucy (2.5) et (2.6), on a:

$$\begin{aligned} Z_{k+1} = & (1 + \varepsilon B_-)(1 - H_- K_-) Z_k + \varepsilon (1 + \varepsilon B_+)(B_+ - B_-) H_+ \hat{x}_k^+ \\ & + (H_+ K_+ (1 + \varepsilon B_+) - H_- K_- (1 + \varepsilon B_-)) \nu_{k+1}^+. \end{aligned} \quad (4.21)$$

Vu que les filtres en question sont à "mémoire courte", d'après l'expression (4.21), l'hypothèse d'indépendance sur A^+ de ν_k^+ et ν_{k+1}^+ , nous permet de supposer "l'indépendance" de Z_k et ν_{k+1}^+ à partir d'un certain instant $i_1 > i_0$.

Remarque 4.2 Sous (HD1), $H_+ K_+ - H_- K_-$ est d'ordre $\mathcal{O}(1)$.

Remarque 4.3 Pour ε "assez petit", on a la majoration

$$0 < (1 + \varepsilon B_-)(1 - H_- K_-) < 1.$$

Pour un intervalle de monotonie suffisamment grand, nous supposons que la situation stationnaire est atteinte. Dans le cas où $B_+ \neq B_-$ nous émettons donc l'hypothèse supplémentaire:

(H4) : $B_+ < 0, B_- < 0$.

On obtient alors

- si $B_+ = B_- \triangleq B$, dans une situation asymptotique,

$$\mathbb{E}[Z^2; A_+] = \varepsilon \frac{(1 + \varepsilon B)^2 (H_+ K_+ - H_- K_-)^2}{1 - (1 + \varepsilon B)^2 (1 - H_- K_-)^2} \vartheta_+.$$

- si $B_+ < 0$; $E[\hat{x}_k^{+2}; A_+]$ admet une solution d'équilibre:

$$E[\hat{x}_k^{+2}; A_+] = \frac{K_+^2}{|B_+|(2 + \varepsilon B_+)} \vartheta^+.$$

En utilisant l'inégalité

$$E[Z_k \hat{x}_k^+; A_+] \leq \frac{1}{2\sqrt{\varepsilon}} E[Z_k^2; A_+] + \frac{\sqrt{\varepsilon}}{2} E[\hat{x}_k^+; A_+],$$

et la majoration suivante, pour ε "assez petit",

$$(1 + \varepsilon B_-)^2 (1 - H_- K_-)^2 + \sqrt{\varepsilon} H_+ |B_+ - B_-| (1 - H_- K_-) (1 + \varepsilon B_+) (1 + \varepsilon B_-) < 1 - H_- K_- (2 - H_- K_-) + c\sqrt{\varepsilon} < 1,$$

on obtient l'estimation

$$\begin{aligned} & E[Z_{k+1}^2; A_+] \\ & \leq (1 - H_- K_- (2 - H_- K_-) + c\sqrt{\varepsilon}) E[Z_k^2; A_+] \\ & \quad + \varepsilon^{3/2} (1 + \varepsilon B_+) (1 + \varepsilon B_-) |B_+ - B_-| (1 - H_- K_-) H_+ E[\hat{x}^{+2}; A_+] \\ & \quad + \varepsilon [H_+ K_+ (1 + \varepsilon B_+) - H_- K_- (1 + \varepsilon B_-)]^2 \vartheta^+ \\ & \quad + \varepsilon^2 (1 + \varepsilon B_+)^2 (B_+ - B_-)^2 H_+^2 E[\hat{x}^{+2}; A_+]. \end{aligned}$$

Dans une situation stationnaire, on a

$$E[Z^2; A_+] \leq \Lambda_+ \varepsilon + c\varepsilon^{3/2},$$

où

$$\Lambda_+ = \frac{(H_+ K_+ - H_- K_-)^2}{H_- K_- (2 - H_- K_-)} \vartheta_+.$$

De plus, pour un certain n_0 , on peut considérer Z_k comme une combinaison linéaire de $\nu_{k-n_0}^+, \dots, \nu_k^+$. Le processus $\xi_k = Z_k \nu_{k+1}^+$ est alors approximativement un processus mélangeant. En conséquence, on approche $\varepsilon \hat{L}_n$ par le processus de diffusion $\{\zeta_t; t = \varepsilon n\}$ solution de l'équation différentielle stochastique

$$d\zeta_t = \Lambda_+/2 dt + \sqrt{\varepsilon \Lambda_+ \vartheta^+} dW_t^+.$$

Comme pour le test de la variation quadratique, on obtient la probabilité d'erreur $p_+ = P(\{\zeta_{T_+^*} = -\varepsilon l_1\} | A_+)$:

$$p_+ = \frac{1 - e^{-\eta_+ l_2}}{e^{\eta_+ l_1} - e^{-\eta_+ l_2}},$$

où T_+^* est le temps d'arrêt défini par

$$T_+^* = \inf\{t : \zeta_t \geq \varepsilon l_2 \text{ ou } \zeta_t \leq -\varepsilon l_1\}$$

et $\eta_+ = 1/\vartheta^+$.

Si on se place dans le cadre de l'hypothèse " A_- ", par un raisonnement identique au précédent, on obtient la probabilité d'erreur $p_- = P(\{\zeta_{T_-} = \varepsilon l_2\} | A_-)$. Elle est donnée par:

$$p_- = \frac{1 - e^{-\eta_- l_1}}{e^{\eta_- l_2} - e^{-\eta_- l_1}},$$

avec $\eta_- = 1/\vartheta_-$.

Pour des valeurs de p_+ et p_- données (⁴), on détermine les bornes de précision l_1 et l_2 et on estime les temps moyens pour prendre une décision:

$$\begin{aligned} E(T_+^*) &= 2\varepsilon \left(\frac{l_2}{\Lambda_+} - p_+ \frac{l_1 + l_2}{\Lambda_+} \right), \\ E(T_-^*) &= 2\varepsilon \left(\frac{l_1}{\Lambda_-} - p_- \frac{l_1 + l_2}{\Lambda_-} \right). \end{aligned}$$

Remarque 4.4 Dans [FJSZ], les auteurs constatent qu'en pratique le temps moyen pour prendre une décision est plus long pour le test du rapport de vraisemblance que pour le test de la variation quadratique.

⁴En général on prend $p_- = p_+ = 0.05$.

5 Sous l'hypothèse (HD2): décider du signe de

x_k

Sous l'hypothèse (HD2): $\Gamma^2 = (H_- \sigma_-)^2 = (H_+ \sigma_+)^2$ et $B_- \neq B_+$, le test de la variation quadratique n'est pas applicable. En remplacement nous considérons un test du type rapport de vraisemblance sur les observations y_k . Puis nous adapterons, sous cette hypothèse, le test du rapport de vraisemblance sur les sorties des filtres de Kalman-Bucy, présenté dans le paragraphe précédent sous l'hypothèse (HD1).

5.1 Test du rapport de vraisemblance sur les accroissements des observations

Ce test est déduit de l'étude du cas $\varepsilon = 0$.

5.1.1 Le cas $\varepsilon = 0$

On considère le système suivant:

$$\begin{cases} x_{k+1} = x_k + \Delta t b(x_k) + \sqrt{\Delta t} \sigma(x_k) u_k, \\ \tilde{y}_k = h(x_k). \end{cases} \quad (5.24)$$

L'équation d'état de ce système généralise celle du système discret (1.2) à un pas de temps Δt .

Le problème est de déterminer le signe de x_k pour $k = i_0, i_0 + 1, \dots, i_0 + m$. Le test d'hypothèses est fait sur les observations $\{\tilde{y}_k, i_0 + 1 \leq k \leq i_0 + m\}$.

On considère la filtration $(\mathcal{F}_{i_0}^k)_k$, où $\mathcal{F}_{i_0}^k = \sigma\{x_{i_0}, u_{i_0}, \dots, u_k\}$, $i_0 \leq k \leq i_0 + m$.

$$\text{Sur } A_+ : \quad \tilde{y}_{k+1} = (1 + \Delta t B_+) \tilde{y}_k + \sqrt{\Delta t} \Gamma u_k^+, \quad (5.24)$$

où $\{u_k^+; k \geq i_0\}$ est un $\mathcal{F}_{i_0}^k$ -bruit blanc gaussien pour une probabilité P^+ telle que $P^+(D) = P(D), \forall D \subset A_+$.

$$\text{Sur } A_- : \quad \tilde{y}_{k+1} = (1 + \Delta t B_-) \tilde{y}_k + \sqrt{\Delta t} \Gamma u_k^-, \quad (5.24)$$

où $\{u_k^-; k \geq i_0\}$ est un $\mathcal{F}_{i_0}^k$ -bruit blanc gaussien pour une probabilité P^- telle que $P^-(D) = P(D), \forall D \subset A_-$.

Le processus $\{\tilde{y}_k; k \geq i_0 + 1\}$ est $\mathcal{F}_{i_0}^{k-1}$ -mesurable (i.e. prévisible) et $\sum_{k=i_0}^{i_0+m} \tilde{y}_k^2 < +\infty$ P -p.s.

On utilise une version discrète du théorème de Girsanov (cf. [Milheiro]) et, vu que $B_- \neq B_+$, on prend pour statistique le logarithme du rapport de vraisemblance:

$$L_n = \ln \frac{dP^+}{dP^-} \Big|_{\mathcal{F}_{i_0+n}},$$

où $1 \leq n \leq m-1$. On a la formule:

$$L_n = \frac{1}{\Gamma^2} \left[(B_+ - B_-) \sum_{k=i_0+1}^{i_0+n} \tilde{y}_k (\tilde{y}_{k+1} - \tilde{y}_k) - \frac{1}{2} \Delta t (B_+^2 - B_-^2) \sum_{k=i_0+1}^{i_0+n} \tilde{y}_k^2 \right]. \quad (5.25)$$

Sur l'événement A_+ , la statistique L_n se réécrit sous la forme:

$$L_n = +\frac{1}{2} \Delta t \frac{(B_+ - B_-)^2}{\Gamma^2} \sum_{k=i_0+1}^{i_0+n} \tilde{y}_k^2 + \sqrt{\Delta t} \frac{B_+ - B_-}{\Gamma} \sum_{k=i_0+1}^{i_0+n} \tilde{y}_k u_k^+. \quad (5.25)$$

Sur l'événement A_- , on a:

$$L_n = -\frac{1}{2} \Delta t \frac{(B_+ - B_-)^2}{\Gamma^2} \sum_{k=i_0+1}^{i_0+n} \tilde{y}_k^2 + \sqrt{\Delta t} \frac{B_+ - B_-}{\Gamma} \sum_{k=i_0+1}^{i_0+n} \tilde{y}_k u_k^-. \quad (5.25)$$

On pose

$$R_n = \Delta t \frac{(B_+ - B_-)^2}{\Gamma^2} \sum_{k=i_0+1}^{i_0+n} \tilde{y}_k^2.$$

On appelle R_n le terme de précision. Pour R_n grand, L_n/R_n est proche de $\pm 1/2$ et la probabilité d'erreur est faible. Si $L_n \geq 0$ on décide A_+ et si $L_n < 0$ on décide A_- .

On établit le résultat suivant:

Théorème 5.1 Soit $r > 0$. $\forall n$ tel que $i_0 < n \leq m-1$, on a:

$$\begin{aligned} P(\{L_n < 0\} \cap \{R_n \geq r\} \cap A_+) &\leq e^{-r/8}, \\ P(\{L_n \geq 0\} \cap \{R_n \geq r\} \cap A_-) &\leq e^{-r/8}. \end{aligned}$$

Preuve La démonstration est identique pour les deux inégalités. Considérons par exemple la première.

On introduit la martingale exponentielle

$$Z_k = \exp \left\{ \lambda \sqrt{\Delta t} \frac{B_+ - B_-}{\Gamma} \sum_{i=i_0+1}^{i_0+k} \tilde{y}_i u_i^+ - \frac{1}{2} \lambda^2 \Delta t \frac{(B_+ - B_-)^2}{\Gamma^2} \sum_{i=i_0+1}^{i_0+k} \tilde{y}_i^2 \right\},$$

avec $\lambda < 0$ et on raisonne comme dans le cas continu (cf. [Roubaud]). ■

Les calculs qui suivent utilisent le fait qu'une situation stationnaire est nécessairement atteinte par $\{\tilde{y}_k\}$, raison pour laquelle nous émettons l'hypothèse supplémentaire:

(H4): $B_+ < 0$ et $B_- < 0$.

On décrit deux procédures de mise en œuvre de ce test sous forme séquentielle.

Première procédure:

Pour r fixé, on introduit le temps d'arrêt N^* :

$$N^* = \inf\{n \geq 1 : R_n \geq r\} \wedge (m+1).$$

- Si $N^* \leq m$:
 - si $L_{N^*} \geq 0$ on décide A_+ ,
 - si $L_{N^*} < 0$ on décide A_- ;
- Sinon on ne peut pas décider.

Nous devons donc proposer une méthode de détermination de la constante r . On considère le problème sur A_+ et sur A_- et on choisit $r = \max(r_-, r_+)$.

Sur l'évènement A_+ , on pose

$$M_n^+ = \sqrt{\Delta t} \frac{B_+ - B_-}{\Gamma} \sum_{k=i_0+1}^{i_0+n} \tilde{y}_k u_k^+.$$

La statistique L_n se réécrit sous la forme:

$$L_n = \frac{1}{2}R_n + M_n^+.$$

On introduit le temps d'arrêt N_+^* :

$$N_+^* = \inf\{n \geq 1 : R_n \geq r_+\} \wedge (m+1).$$

On a que $R_{N_+^*}$ est approximativement égal à r_+ .

On cherche à déterminer une constante r_+ telle que, pour une probabilité $1 - \alpha$ donnée, le terme de précision soit significatif dans le calcul de $L_{N_+^*}$. Dans d'autres termes, on cherche r_+ telle que:

$$P(\{-M_{N_+^*}^+ > \frac{1}{2}r_+\} | A_+) \leq \alpha. \quad (5.26)$$

D'après (5.24), on réécrit \tilde{y}_k sous la forme suivante:

$$\tilde{y}_{k+1} = (1 + \Delta t B_+)^{n_0} \tilde{y}_{k+1-n_0} + \sqrt{\Delta t} \Gamma \sum_{i=0}^{n_0-1} (1 + \Delta t B_+)^i u_{k-i}^+.$$

L'hypothèse (H4) nous permet de considérer \tilde{y}_k comme une combinaison linéaire des variables aléatoires gaussiennes et indépendantes $u_{k-n_0+1}^+, \dots, u_k^+$. Soit $\xi_k = \tilde{y}_k u_k^+$. Le processus $\{\xi_k, i_0 + n_0 - 1 \leq k \leq i_0 + m\}$ est alors approximativement un processus mélangeant. De plus, la valeur stationnaire de $E(\tilde{y}_k^2; A_+)$ est approchée par

$$\Lambda_+ = \frac{\Gamma^2}{-B_+(2 + \Delta t B_+)}.$$

D'après le théorème fonctionnel de limite centrale [Billingsley][th.20.1],

$$\frac{\sqrt{\Delta t}}{\sqrt{\Lambda_+}} \sum_{k=i_0+1}^{i_0+n} \xi_k \approx W_t,$$

où $t = n\Delta t$ et $\{W_t\}$ est un processus de Wiener standard.

On déduit de ce qui précède que M_n^+ suit approximativement une loi normale

$$\mathcal{N}\left(0, n\Delta t \frac{(B_+ - B_-)^2}{-B_+(2 + \Delta t B_+)}\right).$$

Vu que $R_{N_+^*} \approx r_+$, on obtient la relation suivante entre r_+ et N_+^* :

$$r_+ \approx N_+^* \Delta t \frac{(B_+ - B_-)^2}{-B_+(2 + \Delta t B_+)}.$$

Alors l'inégalité (5.26) devient:

$$P\left(\left\{\frac{-M_{N_+^*}^+}{\sqrt{r_+}} > \frac{1}{2}\sqrt{r_+}\right\} | A_+\right) \leq \alpha.$$

On choisit

$$r_+ = 4\lambda^2,$$

avec λ obtenu par la table de la loi normale centrée réduite.

Sur l'évènement A_- , on prend

$$r_- = 4\lambda^2$$

et on a la relation suivante entre r_- et N_-^* :

$$r_- \approx N_-^* \Delta t \frac{(B_+ - B_-)^2}{-B_-(2 + \Delta t B_-)}.$$

Exemple 5.2 Pour $B_- = -1$, $B_+ = -0.25$ et $\alpha = 0.05$, on obtient $r = 10.89$. Pour les estimations des temps moyens, on obtient: $\Delta t N_+^* = 9.66$, $\Delta t N_-^* = 38.52$.

Deuxième procédure:

Il semble que dans le test précédent, le temps d'attente pour prendre une décision soit "trop" long. En effet les bornes r_+ et r_- sont déterminées indépendamment l'une de l'autre. Pour réduire cette durée, nous proposons une autre procédure.

Pour des constantes $l_1 > 0$, $l_2 > 0$ à déterminer, on introduit le temps d'arrêt N^* :

$$N^* = \inf \{n > i_0 \text{ t.q. } L_n \geq l_2 \text{ ou } L_n \leq -l_1\} \wedge m.$$

Règle de décision :

- Si $L_{N^*} \geq l_2$ on décide " A_+ ".
- Si $L_{N^*} \leq -l_1$ on décide " A_- ".
- Sinon, on ne peut pas décider.

On utilise à nouveau l'approximation des processus discrets par des diffusions pour déterminer les constantes l_1 et l_2 .

Sur A_+ , on approche L_n par $\{\zeta_t; t = \Delta t n\}$, où $\{\zeta_t\}$ est le processus de diffusion défini par l'équation différentielle stochastique

$$d\zeta_t = \frac{(B_+ - B_-)^2 \Lambda_+}{2\Gamma^2} dt + \frac{\sqrt{\Lambda_+}(B_+ - B_-)}{\Gamma} dW_t.$$

On introduit le temps d'arrêt T_+^* :

$$T_+^* = \inf \{t \geq 0 : \zeta_t \geq l_2 \text{ ou } \zeta_t \leq -l_1\}.$$

On définit la probabilité $p_+(x) = P_{\zeta_0=x}(\{\zeta_{T_+^*} = -l_1\} | A_+)$. En utilisant la formule de Dynkin, on obtient l'équation différentielle ordinaire suivante:

$$\begin{cases} p_+'' + p_+' = 0 \\ p_+(-l_1) = 1, p_+(l_2) = 0. \end{cases}$$

La résolution de cette équation nous donne la probabilité de "refuser A_+ à tort", $p_+ = p_+(0)$ (cf. [Fleming-Rishel][chapV.th7.1]):

$$p_+ = \frac{1 - e^{-l_2}}{e^{l_1} - e^{-l_2}}.$$

Par un raisonnement identique sur A_- , on obtient l'expression de la probabilité d'erreur $p_- = p_-(0)$:

$$p_- = \frac{1 - e^{-l_1}}{e^{l_2} - e^{-l_1}}.$$

De même, on calcule les temps moyens pour prendre une décision en résolvant deux équations différentielles ordinaires. (cf. [Fleming-Rishel][chapV.th7.1]). Plus précisément $E(T_+^*)$ est donné par la solution de l'équation

$$\begin{cases} \frac{\Lambda_+(B_+ - B_-)^2}{2\Gamma^2} g_+'' + \frac{\Lambda_+(B_+ - B_-)^2}{2\Gamma^2} g_+' + 1 = 0 \\ g_+(-l_1) = g_+(l_2) = 0 \end{cases}$$

au point 0.

On obtient $E(T_-^*)$ de façon similaire:

$$E(T_+^*) = \frac{-2B_+(2 + \Delta t B_+)}{(B_+ - B_-)^2} [l_2 - p_+(l_1 + l_2)],$$

$$E(T_-^*) = \frac{-2B_-(2 + \Delta t B_-)}{(B_+ - B_-)^2} [l_1 - p_-(l_1 + l_2)].$$

Remarque 5.3 Dans les applications, on prend généralement $p_+ = p_- = \alpha$. Ainsi un simple calcul nous donne

$$l_1 = l_2 = \ln \frac{1 - \alpha}{\alpha}.$$

et on note que

$$E(T_-^*) \approx \frac{B_-}{B_+} E(T_+^*).$$

Exemple 5.4 Pour $B_- = -1$, $B_+ = -0.25$ et $\alpha = 0.05$, on obtient $l_1 = l_2 = 2.94$. Les temps moyens sont alors $E(T_+^*) = 4.70$ et $E(T_-^*) = 18.79$.

Remarque 5.5 Pour comparer les temps moyens des deux procédures, on s'intéresse au signe de l'expression:

$$2\lambda^2 - (1 - 2\alpha) \ln \frac{1 - \alpha}{\alpha},$$

avec λ tel que $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\lambda} e^{-x^2/2} dx = \alpha$.

On vérifie numériquement que cette expression est positive, $\forall 0 < \alpha < 1/2$. On en déduit que pour une probabilité d'erreur donnée, les temps moyens de la première procédure sont plus grands que ceux de la deuxième. Cette différence est d'autant plus grande que α est petit.

En préparation de l'étude du cas $\varepsilon > 0$, nous réécrivons la statistique L_n sous une autre forme. En utilisant l'égalité

$$\tilde{y}_k (\tilde{y}_{k+1} - \tilde{y}_k) = -\frac{1}{2} ((\tilde{y}_{k+1} - \tilde{y}_k)^2 - \tilde{y}_{k+1}^2 + \tilde{y}_k^2)$$

on obtient

$$\begin{aligned} L_n &= \frac{B_+ - B_-}{2\Gamma^2} (\tilde{y}_{i_0+n+1}^2 - \tilde{y}_{i_0+1}^2) - \frac{B_+ - B_-}{2\Gamma^2} \sum_{k=i_0+1}^{i_0+n} (\tilde{y}_{k+1} - \tilde{y}_k)^2 \\ &\quad - \Delta t \frac{B_+^2 - B_-^2}{2\Gamma^2} \sum_{k=i_0+1}^{i_0+n} \tilde{y}_k^2. \end{aligned}$$

Lemme 5.6 Sur $A_+ \cup A_-$, on a l'estimation:

$$\sum_{k=i_0+1}^{i_0+n} (\tilde{y}_{k+1} - \tilde{y}_k)^2 - \Gamma^2 n \Delta t = \mathcal{O}(\sqrt{\Delta t}).$$

Preuve Démontrons, par exemple, ce résultat sur l'évènement A_+ .

Sur A_+ , on a :

$$\tilde{y}_{k+1} = \tilde{y}_k + \Delta t H_+ B_+ x_k + \sqrt{\Delta t} \sigma_+ H_+ u_k,$$

d'où l'expression:

$$\begin{aligned} \sum_{k=i_0+1}^{i_0+n} (\tilde{y}_{k+1} - \tilde{y}_k)^2 &= \Delta t^2 H_+^2 B_+^2 \sum_{k=i_0+1}^{i_0+n} x_k^2 + \Delta t \Gamma^2 \sum_{k=i_0+1}^{i_0+n} u_k^2 \\ &\quad + 2 \Delta t^{3/2} H_+^2 \sigma_+ B_+ \sum_{k=i_0+1}^{i_0+n} x_k u_k. \end{aligned}$$

On obtient l'estimation:

$$\mathbb{E} \left[\sum_{k=i_0+1}^{i_0+n} (\tilde{y}_{k+1} - \tilde{y}_k)^2 ; A_+ \right] \leq \Delta t^2 H_+^2 B_+^2 \sum_{k=i_0+1}^{i_0+n} \mathbb{E}(x_k^2) + \Delta t n \Gamma^2.$$

L'égalité

$$\mathbb{E} \left[\left(\sum_{k=i_0+1}^{i_0+n} u_k^2 \right) \left(\sum_{k=i_0+1}^{i_0+n} x_k u_k \right) \right] = 0,$$

entraîne

$$\begin{aligned} & \mathbb{E} \left[\left(\sum_{k=i_0+1}^{i_0+n} (\tilde{y}_{k+1} - \tilde{y}_k)^2 - \Gamma^2 n \Delta t \right)^2 ; A_+ \right] \\ & \leq \mathbb{E} \left[\left(\Delta t^2 H_+^2 B_+^2 \sum_{k=i_0+1}^{i_0+n} x_k^2 + 2 \Delta t^{3/2} H_+^2 \sigma_+ B_+ \sum_{k=i_0+1}^{i_0+n} x_k u_k \right)^2 \right] \\ & \quad + \Delta t^2 \Gamma^4 \mathbb{E} \left[\left(\sum_{k=i_0+1}^{i_0+n} u_k^2 \right)^2 \right] + 2 \Delta t^3 \Gamma^2 H_+^2 B_+^2 \mathbb{E} \left[\left(\sum_{k=i_0+1}^{i_0+n} u_k^2 \right) \left(\sum_{k=i_0+1}^{i_0+n} x_k^2 \right) \right] \\ & \quad - \Delta t^2 n^2 \Gamma^4 - 2 \Delta t^3 n \Gamma^2 H_+^2 B_+^2 \sum_{k=i_0+1}^{i_0+n} \mathbb{E}(x_k^2). \end{aligned} \quad (5.31)$$

On remarque que

$$\mathbb{E} \left[\left(\sum_{k=i_0+1}^{i_0+n} u_k^2 \right)^2 \right] = \sum_{k=i_0+1}^{i_0+n} \mathbb{E}(u_k^4) + n(n-1) = n(n+2),$$

puisque $\{u_k\}$ est un bruit blanc standard. Ainsi on a:

$$\Delta t^2 \Gamma^4 \mathbb{E} \left[\left(\sum_{k=i_0+1}^{i_0+n} u_k^2 \right)^2 \right] - \Delta t^2 n^2 \Gamma^4 = 2 \Delta t^2 n \Gamma^4$$

On utilise l'inégalité $(a+b)^2 \leq 2a^2 + 2b^2$ pour traiter le premier terme de (5.31) et l'inégalité de Cauchy - Schwartz pour le troisième. Vu que les moments d'ordre 2 et 4 de x_k sont finis sur l'intervalle $[a, b]$ et que $n \leq (b-a)/\Delta t$, on obtient la majoration suivante:

$$\mathbb{E} \left[\left(\sum_{k=i_0+1}^{i_0+n} (\tilde{y}_{k+1} - \tilde{y}_k)^2 - \Gamma^2 n \Delta t \right)^2 ; A_+ \right] \leq c_{a,b} \Delta t,$$

où $c_{a,b}$ est une constante positive dépendante de l'intervalle $[a, b]$ fixé.

Sur l'événement A_- , la démonstration est similaire.

■

5.1.2 Le cas $\varepsilon > 0$

On définit la statistique L_n^ε comme il suit:

$$L_n^\varepsilon = \frac{B_+ - B_-}{2\Gamma^2} (y_{i_0+n+1}^2 - y_{i_0+1}^2) - \frac{1}{2} n \varepsilon (B_+ - B_-) - \varepsilon \frac{B_+^2 - B_-^2}{2\Gamma^2} \sum_{k=i_0+1}^{i_0+n} y_k^2. \quad (5.32)$$

On considère $\Delta t = \varepsilon$ et on s'intéresse à l'écart $L_n^\varepsilon - L_n$.

Théorème 5.7 Sur les événements \mathcal{C} et $\hat{\mathcal{C}}$ définis dans le paragraphe 3, on a l'estimation

$$L_n^\varepsilon - L_n = \mathcal{O}(\sqrt{\varepsilon}).$$

Preuve

On démontre d'abord que $L_n^\varepsilon - L_n$ est d'ordre $\mathcal{O}(\sqrt{\varepsilon})$ sur $A_+ \cup A_-$.

$$\begin{aligned} L_n^\varepsilon - L_n &= \frac{B_+ - B_-}{2\Gamma^2} ((y_{i_0+n}^2 - \tilde{y}_{i_0+n}^2) - (y_{i_0+1}^2 - \tilde{y}_{i_0+1}^2)) \\ &\quad + \frac{B_+ - B_-}{2\Gamma^2} \left[\sum_{k=i_0+1}^{i_0+n} (\tilde{y}_{k+1} - \tilde{y}_k)^2 - \Gamma^2 n \varepsilon \right] \\ &\quad - \varepsilon \frac{B_+^2 - B_-^2}{2\Gamma^2} \sum_{k=i_0+1}^{i_0+n} (y_k^2 - \tilde{y}_k^2). \end{aligned}$$

L'égalité

$$y_k^2 - \tilde{y}_k^2 = \sqrt{\varepsilon} v_k (2h(x_k) + \sqrt{\varepsilon} v_k)$$

implique

$$\begin{aligned} L_n^\varepsilon - L_n &= \sqrt{\varepsilon} \frac{B_+ - B_-}{\Gamma^2} (v_{i_0+n+1} h(x_{i_0+n+1}) - v_{i_0+1} h(x_{i_0+1})) \\ &\quad + \frac{B_+ - B_-}{2\Gamma^2} \left[\sum_{k=i_0+1}^{i_0+n} (h(x_{k+1}) - h(x_k))^2 - \Gamma^2 n \varepsilon \right] \\ &\quad + \varepsilon \frac{B_+ - B_-}{2\Gamma^2} (v_{i_0+n+1}^2 + v_{i_0+1}^2) - \varepsilon^{3/2} \frac{B_+^2 - B_-^2}{2\Gamma^2} \sum_{k=i_0+1}^{i_0+n} v_k h(x_k) \\ &\quad + \varepsilon^2 \frac{B_+^2 - B_-^2}{2\Gamma^2} \sum_{k=i_0+1}^{i_0+n} v_k^2. \end{aligned}$$

Sachant que $\{v_k\}$ est un bruit blanc gaussien standard et que tous les moments de x_k sont finis pour $k \leq m$, on déduit à l'aide du lemme 5.6 que $L_n^\varepsilon - L_n$ est d'ordre $\mathcal{O}(\sqrt{\varepsilon})$ sur $A_+ \cup A_-$.

Pour raisonner sur \mathcal{C} , on utilise la majoration

$$\begin{aligned} \mathbf{E}[(L_n^\varepsilon - L_n)^2; \mathcal{C}] &\leq \mathbf{E}[(L_n^\varepsilon - L_n)^2; A_+ \cup A_-] \\ &\quad + \mathbf{E}[(L_n^\varepsilon - L_n)^2; \mathcal{C} \cap (A_+ \cup A_-)^c]. \end{aligned}$$

De l'inégalité de Cauchy-Schwartz, on obtient

$$\mathbf{E}[(L_n^\varepsilon - L_n)^2; \mathcal{C}] \leq \mathbf{E}[(L_n^\varepsilon - L_n)^4]^{1/2} P[\mathcal{C} \cap (A_+ \cup A_-)^c]^{1/2} + C\varepsilon.$$

A l'aide du lemme 3.4, on montre que $\mathbf{E}[(L_n^\varepsilon - L_n)^4] \leq C$ et du théorème 3.1, on déduit le résultat sur l'événement \mathcal{C} .

De même, du théorème 3.6, on déduit le résultat sur l'événement $\hat{\mathcal{C}}$. ■

Etant donné que les bornes déterminées dans le paragraphe 5.1.1 sont d'ordre $\mathcal{O}(1)$, il est donc légitime d'après le résultat précédent de les utiliser pour tester L_n^ε .

D'après la remarque 5.5, nous avons choisi de mettre en œuvre le test selon sa deuxième version.

5.2 Test du rapport de vraisemblance sur les sorties des filtres de Kalman-Bucy

On montre que la statistique \hat{L}_n , utilisée dans le paragraphe 4.2 sous l'hypothèse (HD1), permet encore de décider entre les alternatives " A_+ " et " A_- " sous l'hypothèse (HD2).

Soit $i_0 < i_1 < i_0 + m$. On a

$$\hat{L}_n = \frac{1}{\varepsilon} \sum_{k=i_1}^{i_1+n} Z_k y_{k+1} - \frac{1}{2\varepsilon} \sum_{k=i_1}^{i_1+n} \left((H_+(1 + \varepsilon B_+) \hat{x}_k^+)^2 - (H_-(1 + \varepsilon B_-) \hat{x}_k^-)^2 \right),$$

où $Z_k = H_+(1 + \varepsilon B_+) \hat{x}_k^+ - H_-(1 + \varepsilon B_-) \hat{x}_k^-$.

On rappelle la définition du terme de précision du test:

$$\hat{R}_n = \frac{1}{\varepsilon} \sum_{k=i_1}^{i_1+n} Z_k^2.$$

Sur A_+ , à partir de l'expression récurrente (4.21) de Z_k , on obtient

$$\begin{aligned}
& \mathbf{E}[Z_{k+1}^2; A_+] \\
& \leq (1 + \varepsilon B_-)^2 (1 - H_- K_-)^2 \mathbf{E}[Z_k^2; A_+] \\
& \quad + \varepsilon^2 (1 + \varepsilon B_+)^2 (B_+ - B_-)^2 H_+^2 \mathbf{E}[\hat{x}^{+2}; A_+] \\
& \quad + \varepsilon [H_+ K_+ (1 + \varepsilon B_+) - H_- K_- (1 + \varepsilon B_-)]^2 \vartheta^+ \\
& \quad + 2\varepsilon (1 + \varepsilon B_+) (1 + \varepsilon B_-) |B_+ - B_-| (1 - H_- K_-) H_+ \mathbf{E}[Z_k \hat{x}_k^+; A_+]
\end{aligned}$$

Remarque 5.8 Sous (HD2), la différence $H_+ K_+ - H_- K_-$ est d'ordre $\mathcal{O}(\varepsilon)$, alors que sous (HD1), elle était d'ordre $\mathcal{O}(1)$.

On utilise l'inégalité

$$\begin{aligned}
\mathbf{E}[Z_k \hat{x}_k^+; A_+] & \leq \frac{H_- K_-}{2\varepsilon (1 + \varepsilon B_-)(1 + \varepsilon B_+) |B_+ - B_-| H_+} \mathbf{E}[Z_k^2; A_+] \\
& \quad + \frac{\varepsilon (1 + \varepsilon B_-)(1 + \varepsilon B_+) |B_+ - B_-| H_+}{2 H_- K_-} \mathbf{E}[\hat{x}^{+2}; A_+]
\end{aligned}$$

et les majorations suivantes, pour ε suffisamment petit,

$$\begin{aligned}
0 & < (1 + \varepsilon B_+)^2 (1 - H_- K_-)^2 + H_- K_- (1 - H_- K_-) \\
& < 1 - H_- K_- + C\varepsilon < 1.
\end{aligned}$$

De la même façon que dans le paragraphe 4.2, sous l'hypothèse (H4), on obtient

$$\mathbf{E}[Z^2; A_+] \leq \hat{\Lambda}_+ \varepsilon^2 + c\varepsilon^3,$$

où

$$\hat{\Lambda}_+ = \frac{(B_+ - B_-)^2}{|B_+| (2 + \varepsilon B_+)} \vartheta_+.$$

Sur A_- , on fait un raisonnement similaire.

Remarque 5.9 Le processus $\{Z_k\}$ est d'ordre $\mathcal{O}(\sqrt{\varepsilon})$ sous l'hypothèse (HD1) et d'ordre $\mathcal{O}(\varepsilon)$ sous l'hypothèse (HD2).

De même que pour le test précédent, on considère deux types de procédures.

Première procédure:

Pour r fixé, on calcule \hat{L}_n jusqu'à ce qu'on ait $\hat{R}_n \geq r$. Par un raisonnement similaire à celui du paragraphe 5.1, on choisit $r = \max(r_-, r_+)$ avec

$$r_- = 4\lambda^2\vartheta^- \text{ et } r_+ = 4\lambda^2\vartheta^+.$$

La constante λ est obtenue par inversion de la fonction de répartition de la loi $\mathcal{N}(0, 1)$. Les estimations des temps moyens pour prendre une décision sont données par

$$\begin{aligned} \varepsilon N_-^* &= 4\lambda^2 \frac{-B_-(2 + \varepsilon B_-)}{(B_+ - B_-)^2}, \\ \varepsilon N_+^* &= 4\lambda^2 \frac{-B_+(2 + \varepsilon B_+)}{(B_+ - B_-)^2}. \end{aligned}$$

Remarque 5.10 Les temps moyens pour prendre une décision associés à cette procédure sont identiques à ceux du test précédent quand on considère le même type d'approche (cf. paragraphe 5.1).

Deuxième procédure:

Soit $l_1 > 0$ et $l_2 > 0$ des bornes fixées, on introduit le temps d'arrêt N^* :

$$N^* = \inf\{n > i_1 : \hat{L}_n \geq l_2 \text{ ou } \hat{L}_n \leq -l_1\} \wedge m.$$

Raisonnons sur l'événement A_+ , par exemple.

Nous savons, d'après ce qui précède, que sous l'hypothèse $(H4)$, le processus $\xi_k = Z_k \nu_{k+1}$ est approximativement un processus mélangeant. Par un raisonnement similaire à celui du paragraphe 4.2, pour $t = \varepsilon n$, on approche \hat{L}_n par le processus de diffusion $\{\zeta_t\}$ défini par:

$$d\zeta_t = \hat{\Lambda}_+/2 dt + \sqrt{\vartheta^+ \hat{\Lambda}_+} dW_t^+.$$

On obtient la probabilité d'erreur p_+ , "probabilité de refuser A_+ à tort":

$$p_+ = \frac{1 - e^{-\eta_+ l_2}}{e^{\eta_+ l_1} - e^{-\eta_+ l_2}},$$

où $\eta_+ = 1/\vartheta^+$ ainsi que le temps moyen pour prendre une décision sur A_+ :

$$E(T_+^*) = 2 \left(\frac{l_2}{\hat{\Lambda}_+} - p_+ \frac{(l_1 + l_2)}{\hat{\Lambda}_+} \right).$$

De même, on obtient la probabilité d'erreur p_- , "probabilité de refuser A_- à tort":

$$p_- = \frac{1 - e^{-\eta_- l_1}}{e^{\eta_- l_2} - e^{-\eta_- l_1}},$$

où $\eta_- = 1/\vartheta^-$ et le temps moyen pour prendre une décision sur A_- :

$$E(T_-^*) = 2 \left(\frac{l_1}{\hat{\Lambda}_-} - p_- \frac{(l_1 + l_2)}{\hat{\Lambda}_-} \right).$$

Remarque 5.11 Sous (HD1), les temps d'attente moyens sont d'ordre $\mathcal{O}(\varepsilon)$ tandis que sous (HD2), ils sont d'ordre $\mathcal{O}(1)$. Dans ce deuxième cas, on peut donc s'attendre à ce qu'il y ait des intervalles où aucune décision n'est prise.

Remarque 5.12 Sous l'hypothèse (HD2), l'écart $\vartheta^+ - \vartheta^-$ est d'ordre $\mathcal{O}(\varepsilon)$, donc les bornes l_1 et l_2 sont "très proches". On a une situation similaire à celle décrite dans la remarque 5.5, i.e. les temps moyens d'attente sont plus longs dans la première procédure que dans la deuxième.

Le test du rapport de vraisemblance sur les sorties des filtres de Kalman sera mis en œuvre selon la deuxième procédure.

6 Application

Nous rappelons que notre objectif est d'étudier le comportement des filtres proposés par rapport à la solution optimale du problème de filtrage et de comparer les différentes procédures permettant de les obtenir. Nous devons donc, en premier lieu, définir des critères de comparaison.

6.1 Les critères de comparaison

Une première idée est de considérer, pour une probabilité d'erreur fixée,

- l'amplitude moyenne des intervalles détectés, pour comparer les tests de détection
- le temps moyen nécessaire pour prendre une décision, pour comparer les tests de détermination du signe.

Pour comparer les tests en termes de probabilité d'erreur, il semble naturel de s'intéresser aux pourcentages suivants:

- pourcentage d'intervalles détectés corrects.
- pourcentage de décisions correctes (i.e. nombre de décisions correctes par rapport au nombre de décisions prises).

Nous qualifions une décision de correcte lorsque le signe choisi correspond à celui de la trajectoire sur un intervalle de monotonie correct.

Les temps moyens d'attente pour prendre une décision sont calculés d'abord théoriquement (voir les paragraphes 4 et 5) et sont ensuite confrontés aux résultats obtenus par simulation.

6.2 Exemples

L'étude du problème sous l'hypothèse ($HD2$) étant notre principal centre d'intérêt, nous considérons des exemples vérifiant cette hypothèse et nous appliquons les différents tests décrits dans le paragraphe 5. Nous étudions l'influence de la fonction d'observation sur l'efficacité des tests, dans l'exemple 6.2 et celle de la dérive dans l'exemple 6.3. Nous nous intéressons ensuite à un exemple vérifiant l'hypothèse ($HD1$), présenté dans [FJSZ]. Ceci nous permettra d'illustrer que dans ce cas, les décisions sont prises beaucoup plus rapidement que sous ($HD2$). Nous justifierons ce fait par le comportement de la solution de l'équation de Zakai.

Dans le paragraphe 5, sous (HD2), nous avons proposé une méthode de détermination des bornes pour les tests du signe et d'estimation des temps d'attente sous l'hypothèse (H4): $B_- < 0$ et $B_+ < 0$. Il est intéressant d'étudier les résultats obtenus dans le cas où l'hypothèse (H4) n'est pas vérifiée. C'est ce que nous nous proposons de faire dans l'exemple 6.6.

Pour chaque exemple, nous commentons les résultats statistiques sur le comportement des tests. On peut consulter ces résultats dans l'annexe A.

Remarque 6.1 On rappelle que la détermination des bornes utilisées dans les tests de décision sur le signe repose sur la possibilité d'atteindre une situation stationnaire. Après la détection d'un passage à zéro, nous devons attendre un "certain" temps, n_0 , avant de cumuler les statistiques de décision des tests. Nous avons choisi pour ces exemples $n_0 = 6$.

6.2.1 Exemples sous l'hypothèse (HD2)

Exemple 6.1 On considère le système

$$\begin{cases} x_{k+1} = x_k + 0.01 b(x_k) + \sqrt{0.01} u_k, & x_0 \sim \mathcal{N}(-5, 0.1) \\ y_k = |x_k| + \sqrt{0.01} v_k, \end{cases}$$

avec

$$b(x) = \begin{cases} -x, & x \leq 0 \\ -0.25 x, & x \geq 0. \end{cases}$$

On s'intéresse tout d'abord à une trajectoire simulée sur l'intervalle de temps $[0, 10]$.

Dans ce cas, les résultats des différentes procédures étant similaires, nous illustrons par la figure 1, uniquement ceux des tests de détection et du rapport de vraisemblance sur les observations.

Dans cette figure, la trajectoire du signal à estimer est tracée en trait plein et le filtre proposé en pointillé. Les intervalles de monotonie détectés sont signalés en noir dans le cas où une décision est prise sur le signe et en gris s'il n'y a pas de décision. Nous rappelons que sur un intervalle de monotonie, si le test appliqué décide "signe positif", (respectivement "signe négatif") on trace la sortie du filtre de Kalman-Bucy (2.5) (respectivement la sortie du filtre (2.6)). Si le test ne prend pas de décision, les sorties des deux filtres sont représentées.

Il est intéressant de regarder le comportement de la solution optimale pour ce problème. Dans la figure 2 on représente, en plus de la trajectoire (en trait plein), l'estimation optimale obtenue par résolution numérique de l'équation de Zakai (en trait pointillé), entourée par sa région de confiance correspondante à

une probabilité de 95%. Un générateur automatique de programmes FORTRAN a été utilisé pour obtenir cette solution.

Dans les figures 3 à 6, on trace la densité conditionnelle à divers instants de l'intervalle $[0, 10]$. Le trait vertical correspond à la valeur du signal simulé à l'instant t . On note qu'au moment où la décision sur le signe est prise, la loi conditionnelle approche une loi gaussienne et que près d'un passage à zéro, la loi conditionnelle a deux "bosses".

On peut constater, par exemple, qu'à l'instant $t = 6.5$ les tests ayant déjà détecté un intervalle de monotonie, ne sont pas encore capables de décider sur le signe. Cette décision sera prise aux environs de $t = 8.48$. En regardant le comportement de la solution de l'équation de Zakai, on remarque l'existence de deux "bosses" à partir de l'instant $t = 3$, approximativement. A l'instant $t = 8.48$, instant où les tests de décision sur le signe sont capables de prendre une décision, une des "bosses" de la densité conditionnelle est déjà négligeable (voir la figure 5), ce qui explique le fait qu'une décision puisse être prise. A cet instant, le filtre proposé et l'estimation optimale ont un comportement similaire.

On s'intéresse maintenant aux résultats numériques obtenus pour une trajectoire simulée sur l'intervalle de temps $[0, 100]$ (voir les tableaux 1 et 2).

On constate que

- Le test de détection sur les sorties du filtre de Kalman-Bucy permet de détecter des intervalles de monotonie plus longs que le test sur les observations; cependant l'erreur commise est aussi plus grande.
- Le fait d'augmenter la borne de détection (par exemple en diminuant la probabilité d'erreur α) entraîne évidemment la détection d'intervalles de monotonie plus petits. Il semble que nous contrôlons moins bien la probabilité d'erreur du test sur la sortie des filtres de Kalman-Bucy que celle du test sur les observations. Voyons par exemple ce qui se passe quand on prend $\alpha_d = 5\%$ (voir le tableau 2). Le pourcentage d'erreur de détection p_2 est supérieur à celui qui avait été imposé.
- Les résultats des tests de décision sur le signe basés sur le rapport de vraisemblance sur les observations ou sur les sorties du filtre de Kalman sont comparables. Toutes fois, le temps moyen pour prendre une décision semble être légèrement plus petit dans le premier cas que dans le deuxième.
- Le fait de diminuer la probabilité d'erreur pour la décision sur le signe entraîne naturellement une réduction du pourcentage d'intervalles avec décision et des temps d'attente plus grands pour prendre ces décisions.

Remarque 6.1 Les temps moyens pour prendre une décision calculés théoriquement sont beaucoup plus élevés que ceux vérifiés en pratique.

Exemple 6.2 On modifie la fonction d'observation dans l'exemple 6.1. On considère:

$$h(x) = 0.4 |x|.$$

On obtient les résultats présentés dans le tableau 3.

Le fait de diminuer le coefficient $|H_+| = |H_-|$ semble avoir pour effet l'augmentation des pourcentages de détection d'intervalles corrects. Autrement dit, lorsqu'on diminue le coefficient $|H_+| = |H_-|$, les constantes des tests de détection augmentent. On détecte plus de passages à zéro et les longueurs des intervalles de monotonie diminuent. Ce comportement des tests est satisfaisant vu qu'une diminution du coefficient $|H_+| = |H_-|$ est équivalente à une augmentation du bruit d'observation.

On vérifie également une "légère" diminution des temps d'attente pour prendre une décision. Ceci s'explique du fait que les statistiques des tests du signe sont calculés loin des passages à zéro.

Dans ce cas, le test de détection des passages à zéro basé sur les sorties d'un des filtres de Kalman-Bucy est meilleur que celui basé sur les observations.

Exemple 6.3 On considère les mêmes paramètres que dans l'exemple précédent sauf pour la dérivée $b(x)$:

$$b(x) = \begin{cases} -5x, & x \leq 0 \\ -0.25x, & x \geq 0. \end{cases}$$

Les résultats sont à consulter dans le tableau 4.

On remarque que le pourcentage de décisions est maintenant plus élevé et que le temps moyen pour prendre une décision diminue considérablement. Ceci illustre parfaitement les calculs théoriques présentés dans le paragraphe 5. Le fait de diminuer le rapport $|B_+|/(B_+ - B_-)^2$ (respectivement $|B_-|/(B_+ - B_-)^2$) a pour effet d'augmenter le terme de précision dans les statistiques L et \hat{L} . On obtient donc de meilleurs résultats, notamment pour le temps d'attente "côté positif" (respectivement "côté négatif").

6.2.2 Exemple sous l'hypothèse (HD1)

Exemple 6.4 Nous changeons la fonction d'observation de l'exemple 6.1:

$$h(x) = \begin{cases} -x, & x \leq 0, \\ 2x, & x \geq 0. \end{cases}$$

Le système considéré vérifie l'hypothèse (HD1).

Par les figures 7 à 10, nous illustrons les résultats des différentes procédures appliquées à une trajectoire simulée sur l'intervalle de temps $[0, 10]$. Nous remarquons que le test de la variation quadratique décide plus souvent que le test sur les sorties des filtres de Kalman-Bucy.

Le comportement de l'estimation optimale et sa région de confiance à 95% représentés dans la figure 11, montrent clairement que dans ce cas, la loi conditionnelle approche beaucoup plus rapidement une loi gaussienne après un passage à zéro que dans l'exemple 6.1 (voir la figure 2). Ceci justifie le fait que nous décidons du signe de x_k plus souvent sous l'hypothèse (HD1) que sous (HD2).

La comparaison entre les figures 12 à 14 et les figures 3 à 6 (voir l'exemple 6.1) illustre à nouveau la remarque précédente. Dans ce cas, une des deux "bosses" de la loi conditionnelle s'estompe rapidement aux environs des instants de décision ($t = 6.86$ et $t = 6.88$), tandis que dans l'exemple 6.1 on ne retrouve un tel comportement qu'aux alentours de l'instant $t = 10$.

Les résultats numériques des tests appliqués à une trajectoire simulée sur l'intervalle de temps $[0, 100]$ (voir le tableau 5) mettent également en évidence que les temps d'attente pour prendre une décision sous l'hypothèse (HD1) sont nettement plus courts que sous l'hypothèse (HD2).

Remarque 6.5 Quelques remarques sur les résultats des procédures appliquées à plusieurs trajectoires simulées sur un intervalles de temps $[0, 100]$.

- Les tests détectent moins de passages à zéro lorsque l'écart $|H_+ - H_-|$ augmente et le taux d'erreur est supérieur au taux souhaité. De façon plus nette que sous (HD2), nous contrôlons "moins bien" la probabilité d'erreur du test de détection sur la sortie d'un des filtres de Kalman-Bucy que celle du test sur les observations.
- Pour une différence $|\sigma_+ H_+ - \sigma_- H_-| = 0.5$, les tests de détermination du signe ont des taux d'erreur supérieurs à ceux demandés. Le test de la variation quadratique a tendance à décider plus souvent que le test du rapport de vraisemblance sur les sorties des filtres de Kalman-Bucy, mais semble généralement moins fiable. (cf. [FJSZ]).

- En augmentant l'écart $|\sigma_+ H_+ - \sigma_- H_-|$, les tests de détermination du signe décident plus souvent et prennent moins de fausses décisions.

6.2.3 Un exemple sous l'hypothèse (HD2) ne vérifiant pas (H4)

Exemple 6.6 On considère le système suivant:

$$\begin{cases} x_{k+1} = x_k + 0.01 b(x_k) + \sqrt{0.01} u_k, & x_0 \sim \mathcal{N}(-5, 0.1) \\ y_k = |x_k| + \sqrt{0.01} v_k, \end{cases}$$

avec

$$b(x) = \begin{cases} -x & x \leq 0 \\ 0.2x & x \geq 0 \end{cases}$$

Remarque 6.7 Le coefficient de dérive B_+ étant positif et B_- étant négatif, il y a “peu” de passages à zéro et le signal a tendance à “fuir” vers l'infini dès qu'il est du côté positif.

Nous étudions un cas où l'hypothèse (HD2) est vérifiée mais pas (H4). Cependant on applique les tests présentés dans la section 5. On rappelle que les bornes utilisées dans ces tests ont été déterminées sous l'hypothèse (H4) et ne sont plus justifiées dans cet exemple. Il est donc intéressant de regarder les résultats qu'elles engendrent dans une telle situation.

Dans la figure 15, nous traçons le filtre approché construit à partir du test de détection et du test de rapport de vraisemblance basés sur les observations. Les résultats obtenus par les autres procédures sont similaires.

En comparant le filtre proposé et l'estimation optimale (voir la figure 16), on remarque que leurs comportements sont semblables sur l'intervalle de monotonie détecté $[0, 3]$ et également à partir de l'instant $t = 7$, environ. Ceci est satisfaisant vu que sur l'intervalle de monotonie $[5.80, 10]$, les instants de décision sur le signe sont $t = 7.15$, pour le test sur les observations, et $t = 7.20$ pour le test sur les sorties des filtres de Kalman-Bucy.

Ainsi, pour certains cas ne vérifiant pas l'hypothèse (H4), on peut encore utiliser les bornes calculées sous cette hypothèse et obtenir des résultats “corrects”.

Remarque 6.8 Dans cette situation où l'hypothèse (H4) n'est pas vérifiée, nous ne savons pas estimer théoriquement le temps d'attente d'une décision.

7 Conclusion

On a décrit des tests permettant de déterminer les intervalles de linéarité de la fonction d'observation, pour un système linéaire par morceaux en temps discret avec petit bruit d'observation. Deux types de tests sont nécessaires: un premier test permet de détecter les intervalles où il n'y a pas de passages à zéro, un deuxième test permet de décider du signe pris par la trajectoire sur cet intervalle. Sous l'hypothèse (*HD2*), différents tests ont été proposés. Pour la détection des intervalles de monotonie, on a le choix entre un test basé sur les observations et un test basé sur la sortie d'un des filtres de Kalman-Bucy. Pour l'étape de décision du signe, on a décrit deux tests du type rapport de vraisemblance; l'un basé sur les observations et l'autre sur les sorties des filtres de Kalman-Bucy.

L'efficacité des tests de détection dépend de la méthode de détermination de leur borne. Pour le test basé sur les observations, une formule permettant d'obtenir cette valeur a été justifiée théoriquement. Ceci n'a pas été de même pour le test basé sur la sortie d'un des filtres de Kalman-Bucy. Il nous semble donc naturel que ce deuxième test se comporte moins bien que le premier. Quant aux tests de détermination du signe, on a vérifié que leurs comportements sont semblables. Des formules permettant de déterminer les bornes et les temps moyens d'attente pour prendre une décision ont été proposées sous l'hypothèse (*H4*). On remarque de manière générale, que les temps moyens d'attente calculés théoriquement sont plus longs que ceux obtenus en pratique. On note également que l'ordre de grandeur des temps moyens d'attente théoriques et empiriques est plus grand sous l'hypothèse (*HD2*) que sous (*HD1*).

On peut envisager d'étendre cette étude à des systèmes où la fonction d'observation h possède plusieurs points critiques dans le cas où l'allure de cette fonction permet d'appliquer les procédures décrites localement. On peut aussi considérer des systèmes de dimension $n \geq 1$, avec des bruits corrélés.

Une extension au cas de fonctions monotones par morceaux existe dans [Fleming-Pardoux], sous l'hypothèse (*HD1*). Des essais ont été faits par [FJSZ] dans le cas d'une dynamique non linéaire. Il reste à justifier son analogue sous l'hypothèse (*HD2*).

Références

- [Bensoussan] A. Bensoussan : *On some approximation techniques in non-linear filtering*, Stoch. Diff. Syst., Stoch. Control and Appl., W.H. Fleming & P.L. Lions eds., IMA Vol 10, Springer, 1987.
- [Billingsley] P. Billingsley : *Convergence of Probability Measures*, John Wiley & Sons, New York, 1968.
- [FJP] W.H. Fleming, D. Ji and E. Pardoux : *Piecewise linear filtering with small observation noise*, Proc. 8th Int. Conference on Analysis and Optimization of Systems, Antibes 1988, Lect. Notes in Control and Info. Sci., n° 111, (pp 725-739), Springer, 1988.
- [FJSZ] W.H. Fleming, D. Ji, P. Salame and Q. Zhang : *Discrete Time Piecewise Linear Filtering with Small Observation Noise*, Brown University Providence RI 02912, Division of Applied Mathematics, September 1988 LCDS/CCS 88-27.
- [Fleming-Pardoux] W.H. Fleming, E. Pardoux : *Piecewise monotone filtering with small observation noise*, Siam J. Control, Vol. 27, No. 5, pp 1156-1181, Sep. 1989.
- [Fleming-Rishel] W. H. Fleming and R.W. Rishel : *Deterministic and Stochastic Optimal Control*, Tome 1, Applications of Mathematics, Springer - Verlag, 1975.
- [Ji] D. Ji : *Nonlinear filtering with small observation noise*, Ph. D. Thesis, Brown University, 1987.
- [KBS] R. Katzur, B.Z. Bobrovsky and Z. Schuss : *Asymptotic analysis of the optimal filtering problem for one-dimensional diffusions measured in a low noise channel*, SIAM J. Appl. Math. 44 (1984), Part I: 591-604, Part II: 1176-1191.
- [Le Gland] F. Le Gland : *Estimation de paramètres dans les processus stochastiques, en observation incomplète - Application à un problème de radio-astronomie*, Thèse de Docteur-Ingenieur, Université Paris IX-Dauphine, Nov. 1981.

- [Milheiro] P. Milheiro de Oliveira : *Filtres approchés pour un problème de filtrage non linéaire discret avec petit bruit d'observation*, RR 1142, INRIA, 1989.
- [Picard] J. Picard : *Nonlinear filtering of one-dimensional diffusions in the case of a high signal-to-noise ratio*. 1986 - SIAM J. Appl. Math. (46),6,1098-1125.
- [Picard] J. Picard : *Nonlinear filtering and smoothing with high signal-to-noise ratio*, Stochastic Processes in Physics and Engineering (Bielefeld 1986), Reidel 1988.
- [Picard] J. Picard : *Efficiency of the extended Kalman filter for nonlinear systems with small observation noise*, Rapport de Recherche INRIA n° 1068, 1989.
- [Roubaud] M.C. Roubaud : *Filtrage linéaire par morceaux avec petit bruit d'observation*, Thèse, Université de Provence, Nov. 1990.

A Résultats numériques et représentations graphiques des exemples présentés dans la section 6

A.1 Résultats numériques

Nous introduisons les notations suivantes:

TDO: test de détection sur les observations.

TDK: test de détection sur la sortie d'un des filtres de Kalman-Bucy.

TLO: test du rapport de vraisemblance sur les observations.

TLK: test du rapport de vraisemblance sur les sorties des filtres de Kalman-Bucy.

α_d : niveau d'erreur pour le test de détection.

α_s : niveau d'erreur pour le test de détermination du signe.

C : borne du test de détection.

$-l_1, l_2$: intervalle de confiance du test de détermination du signe.

ET_-^*, ET_+^* : temps moyens théoriques d'attente d'une décision.

p_1 : pourcentage de pas de temps détectés.

p_2 : pourcentage d'intervalles de monotonie corrects.

p_3 : pourcentage de pas de temps avec décision.

p_4 : pourcentage de décisions correctes.

\bar{T}_-, \bar{T}_+ : moyennes empiriques des temps d'attente d'une décision.

Test de détection	TDO	TDO	TDK	TDK
Test du signe	TLO	TLK	TLO	TLK
C	0.143	0.143	0.143	0.143
$-l_1, l_2$	-2.94 2.94	-7.70 7.67	-2.94 2.94	-7.70 7.67
ET_-^*, ET_+^*	18.8 4.71	18.9 4.69	18.8 4.71	18.9 4.69
p_1	84.5%	84.5%	86.3%	86.3%
p_2	96.3%	96.3%	93.2%	93.2%
p_3	40.6%	40.6%	41.7%	41.7%
p_4	100%	100%	100%	100%
\bar{T}_-, \bar{T}_+	1.3 1.5	1.3 1.5	2.1 1.6	1.7 1.6

Tableau 1: résultats numériques pour l'exemple 6.1 ($\alpha_d = 5\%$, $\alpha_s = 5\%$)

Test de détection	TDO	TDO	TDK	TDK
Test du signe	TLO	TLK	TLO	TLK
C	0.17	0.17	0.17	0.17
$-l_1 \ l_2$	-2.94 2.94	-7.70 7.67	-2.94 2.94	-7.70 7.67
$ET_-^* \ ET_+^*$	18.8 4.71	18.9 4.69	18.8 4.71	18.9 4.69
p_1	82.4%	82.4%	84.2%	84.2%
p_2	97.9%	97.9%	97.5%	97.5%
p_3	39.0%	39.0%	41.4%	41.4%
p_4	100%	100%	100%	100%
$\bar{T}_- \ \bar{T}_+$	1.3 1.5	1.3 1.5	2.1 1.5	1.7 1.5

Tableau 2: résultats numériques pour l'exemple 6.1 ($\alpha_d = 2.5\%$, $\alpha_s = 5\%$)

Test de détection	TDO	TDO	TDK	TDK
Test du signe	TLO	TLK	TLO	TLK
C	0.121	0.121	0.302	0.302
$-l_1 \ l_2$	-2.94 2.94	-4.38 4.36	-2.94 2.94	-4.38 4.36
$ET_-^* \ ET_+^*$	18.8 4.71	18.9 4.70	18.8 4.71	18.9 4.70
p_1	70.8%	70.8%	75.4%	75.4%
p_2	98.4%	98.4%	98.8%	98.8%
p_3	35.0%	34.9%	40.1%	38.8%
p_4	100%	100%	100%	100%
$\bar{T}_- \ \bar{T}_+$	0.0 1.1	0.05 1.2	0.0 1.4	0.05 1.5

Tableau 3: résultats numériques pour l'exemple 6.2 ($\alpha_d = 5\%$, $\alpha_s = 5\%$)

Test de détection	TDO	TDO	TDK	TDK
Test du signe	TLO	TLK	TLO	TLK
C	0.144	0.144	0.144	0.144
$-l_1 \ l_2$	-2.94 2.94	-7.69 7.54	-2.94 2.94	-7.69 7.54
$ET_-^* \ ET_+^*$	2.35 0.117	2.40 0.115	2.35 0.117	2.40 0.115
p_1	85.5%	85.5%	87.0%	87.0%
p_2	95.3%	95.3%	92.5%	92.5%
p_3	66.8%	66.8%	69.3%	68.4%
p_4	95.8%	100.0%	90.0%	94.7%
$\bar{T}_- \ \bar{T}_+$	0.0 0.28	0.12 0.29	0.0 0.30	0.12 0.33

Tableau 4: résultats numériques pour l'exemple 6.3 ($\alpha_d = 5\%$, $\alpha_s = 5\%$)

Test de détection	TDO	TDO	TDK	TDK
Test du signe	TLO	TLK	TLO	TLK
C	0.165	0.165	0.165	0.165
$-l_1 \ l_2$	-4.70 2.37	-15.80 7.80	-4.70 2.37	-15.80 7.80
$ET_-^* \ ET_+^*$	0.44 0.13	2.46 0.44	0.44 0.13	2.46 0.44
p_1	86.6%	86.6%	88.2%	88.2%
p_2	93.5%	93.5%	87.5%	87.5%
p_3	71.2%	63.4%	76.1%	65.4%
p_4	82.4%	87.0%	71.1%	66.7%
$\bar{T}_- \ \bar{T}_+$	0.38 0.29	1.60 0.28	0.40 0.37	1.60 0.30

Tableau 5: résultats numériques pour l'exemple 6.4 ($\alpha_d = 5\%$, $\alpha_s = 5\%$)

A.2 Représentations graphiques

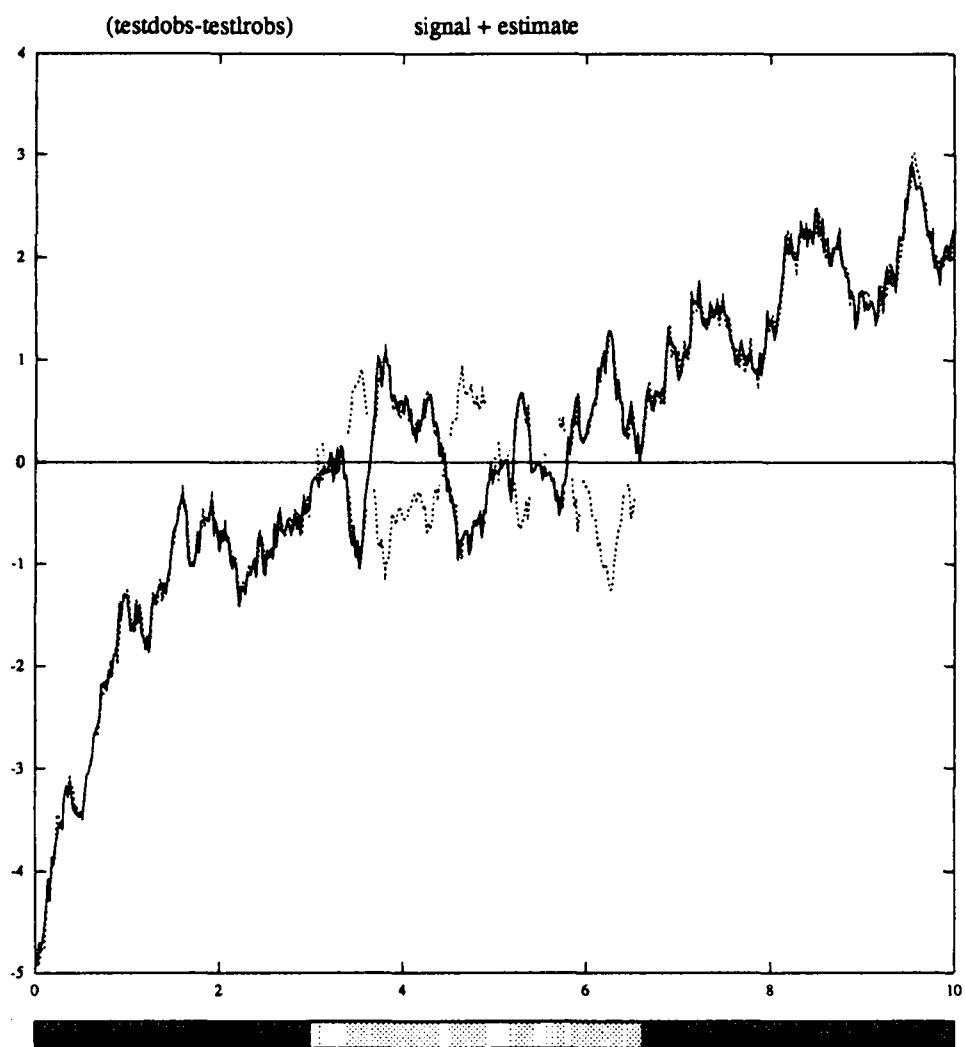


Figure 1: Sous ($HD2$), exemple 6.1: tests de détection des passages à zero et de décision du signe basés sur les observations.

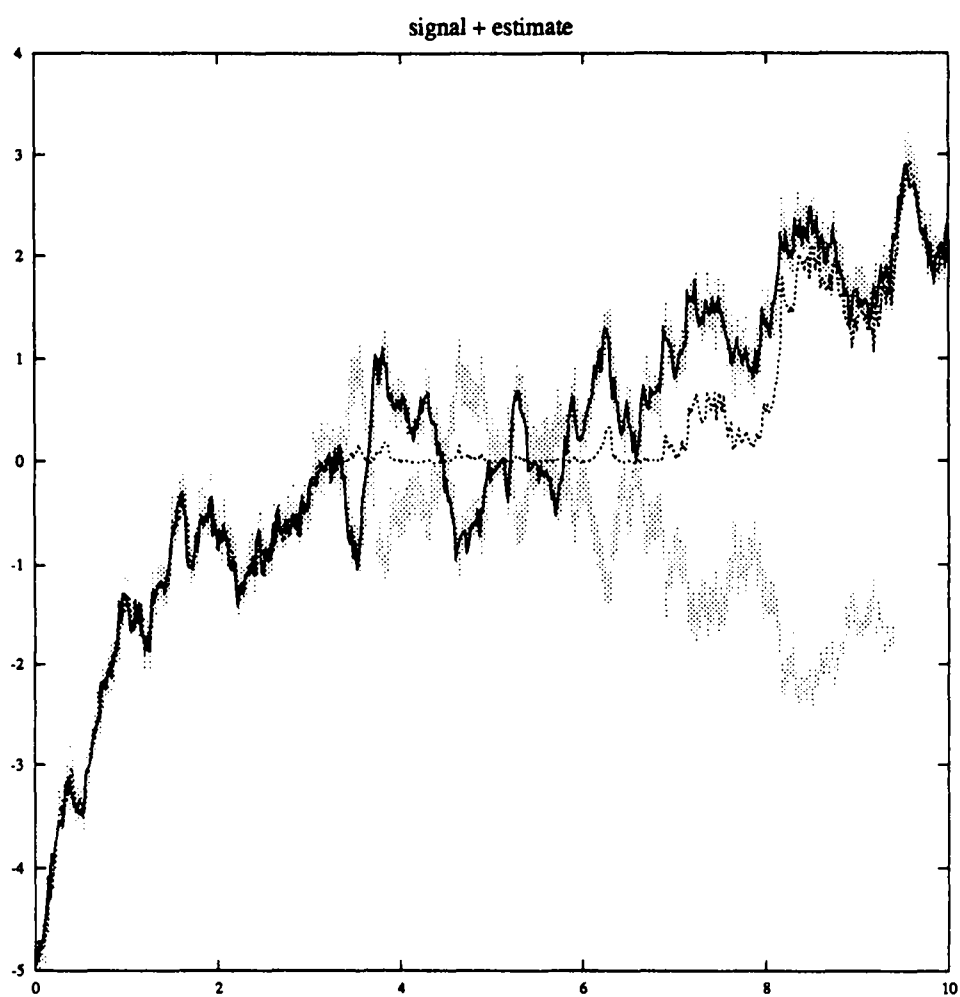


Figure 2: Sous ($HD2$), exemple 6.1: solution de l'équation de Zakai (le niveau de gris correspond à une région de confiance à 95%).

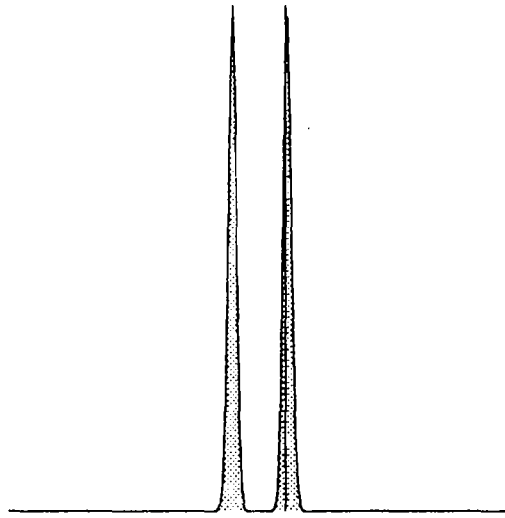


Figure 3: Sous ($HD2$), exemple 6.1: densité conditionnelle à l'instant $t = 6.5$

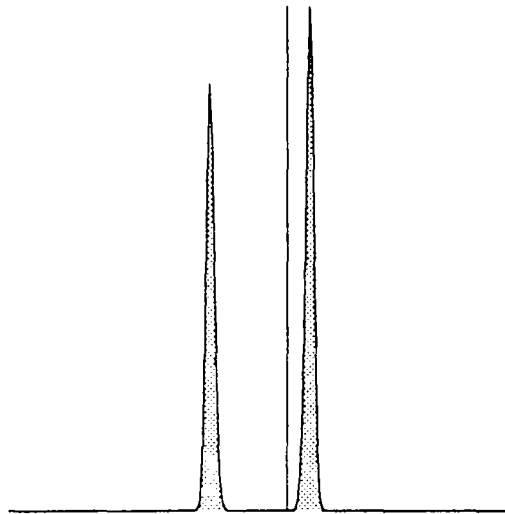


Figure 4: Sous ($HD2$), exemple 6.1: densité conditionnelle à l'instant $t = 7$

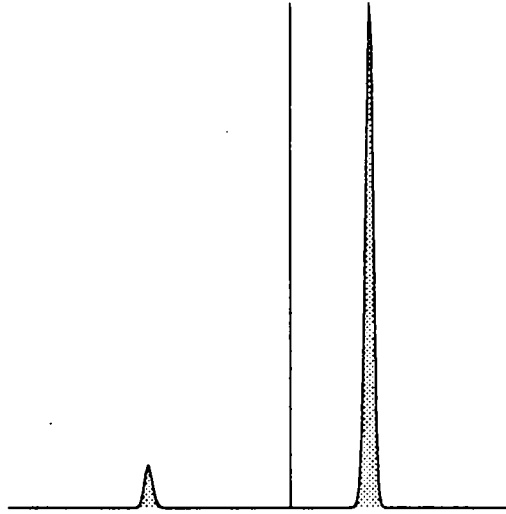


Figure 5: Sous ($HD2$), exemple 6.1: densité conditionnelle à l'instant $t = 8.48$

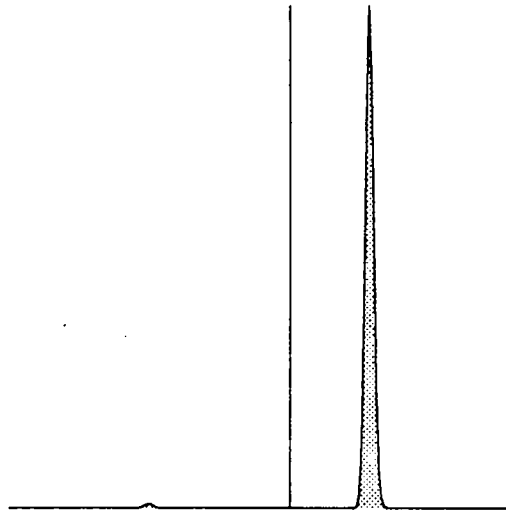


Figure 6: Sous ($HD2$), exemple 6.1: densité conditionnelle à l'instant $t = 10$

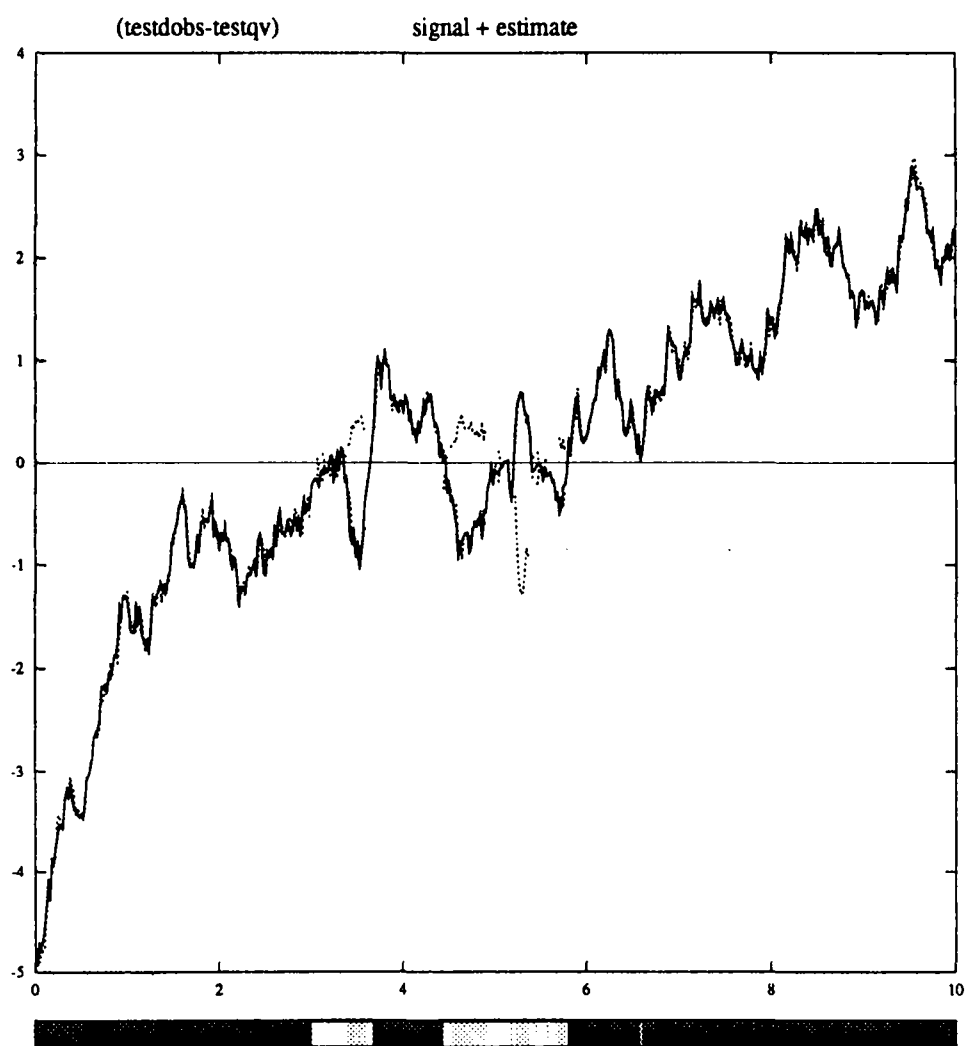


Figure 7: Sous ($HD1$), exemple 6.4: test de détection des passages à zero basé sur les observations et test de décision du signe basé sur la variation quadratique.

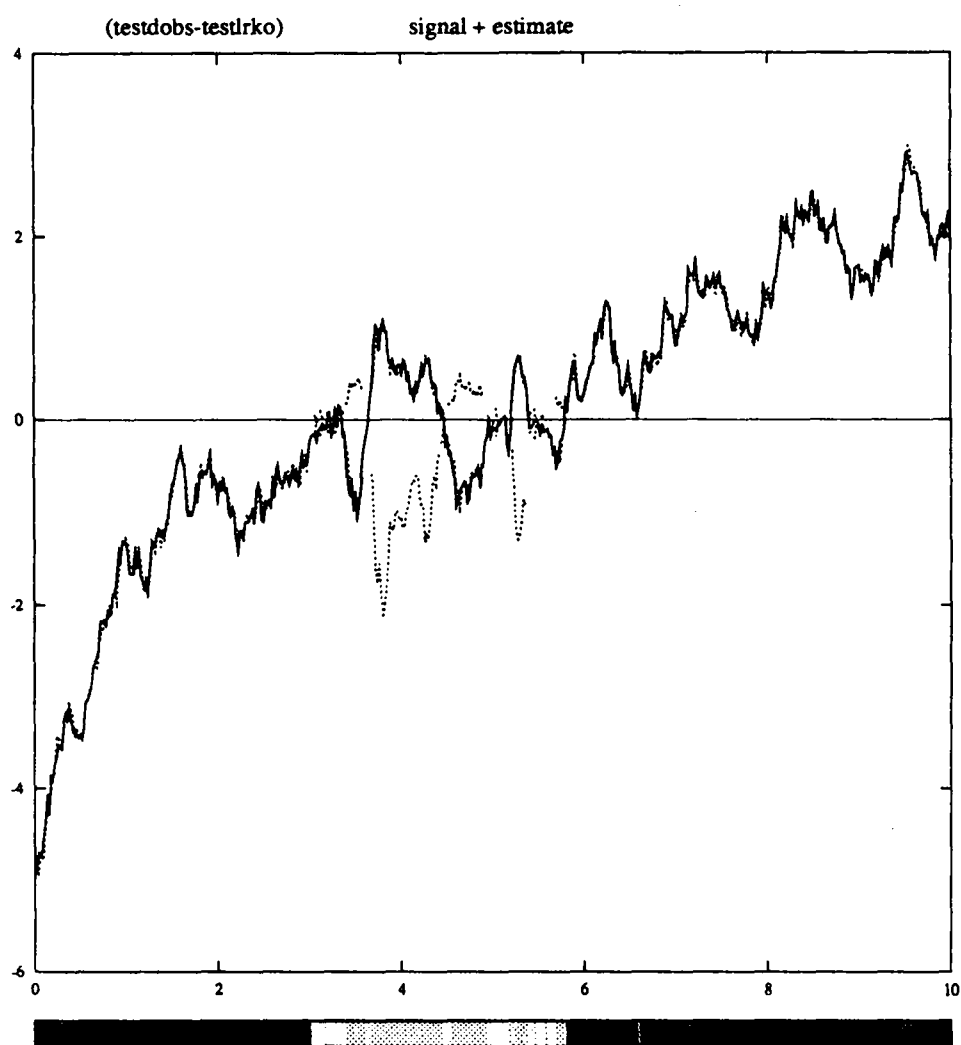


Figure 8: Sous ($HD1$), exemple 6.4: test de détection des passages à zero basé sur les observations et test de décision du signe basé sur la sortie des filtres de Kalman-Bucy.

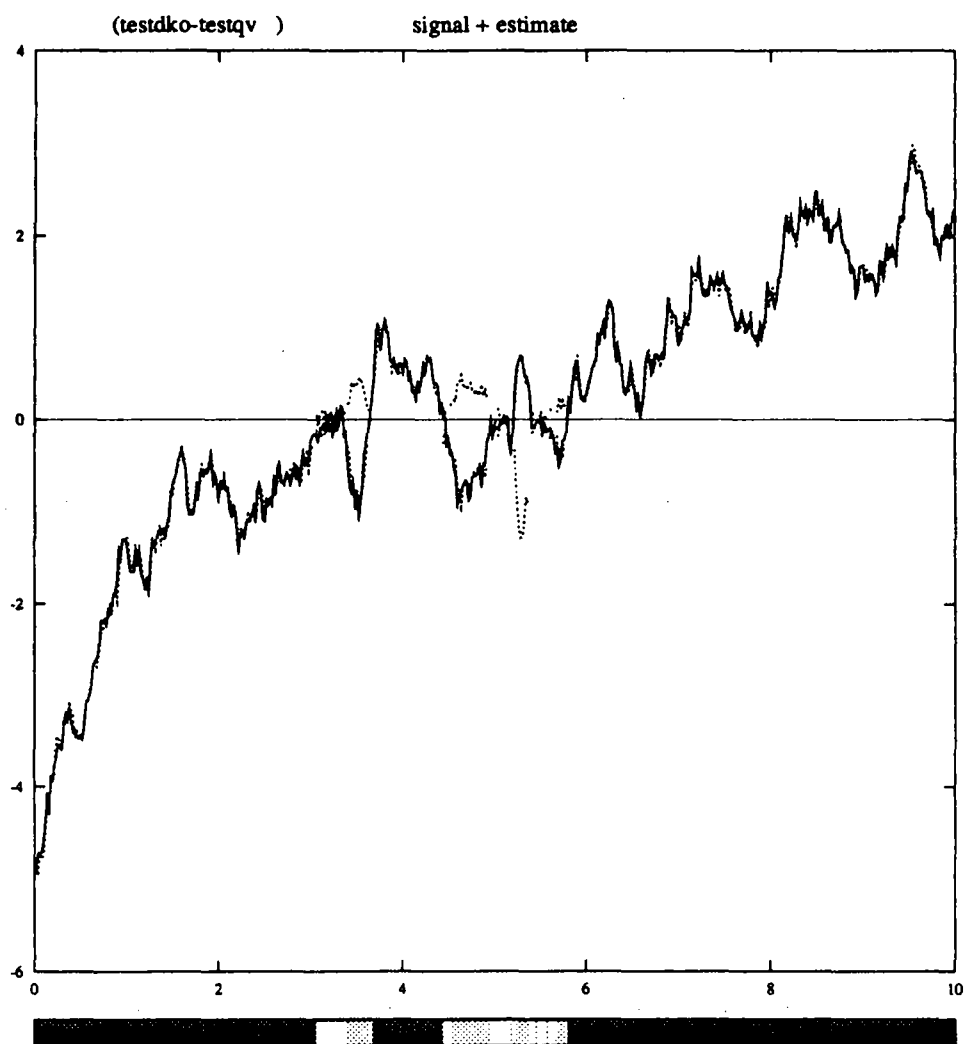


Figure 9: Sous ($HD1$), exemple 6.4: test de détection des passages à zero basé sur la sortie des filtres de Kalman-Bucy et test de décision du signe basé sur la variation quadratique.

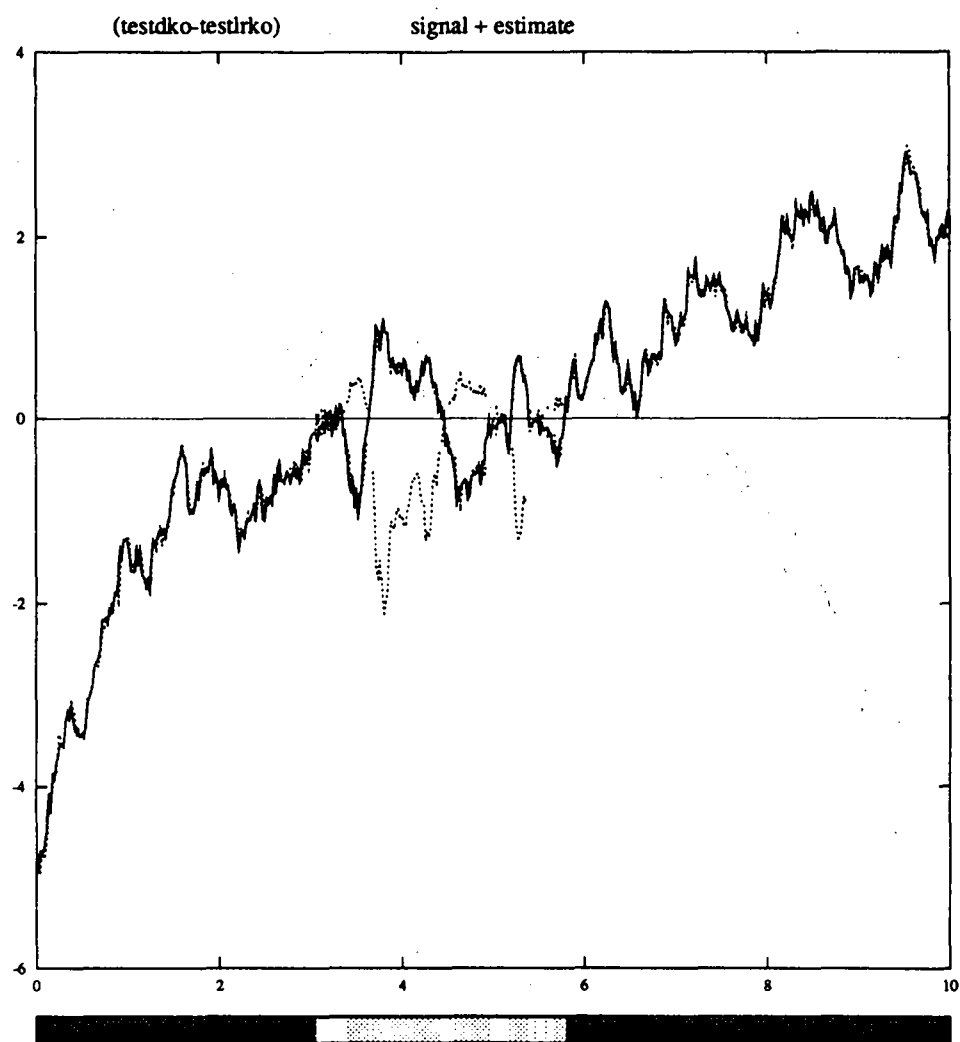


Figure 10: Sous ($HD1$), exemple 6.4: tests de détection des passages à zero et de décision du signe basés sur la sortie des filtres de Kalman-Bucy.

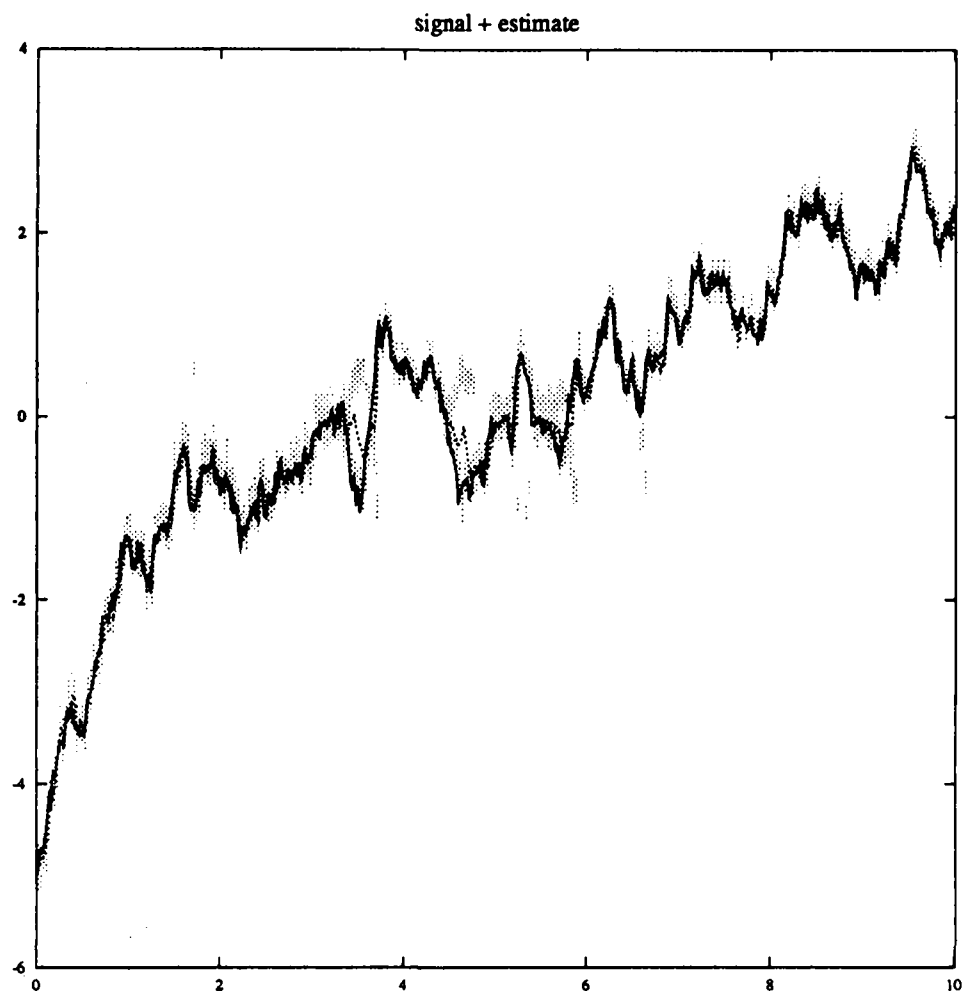


Figure 11: Sous $(HD1)$, exemple 6.4: solution de l'équation de Zakai (le niveau de gris correspond à une région de confiance à 95%).

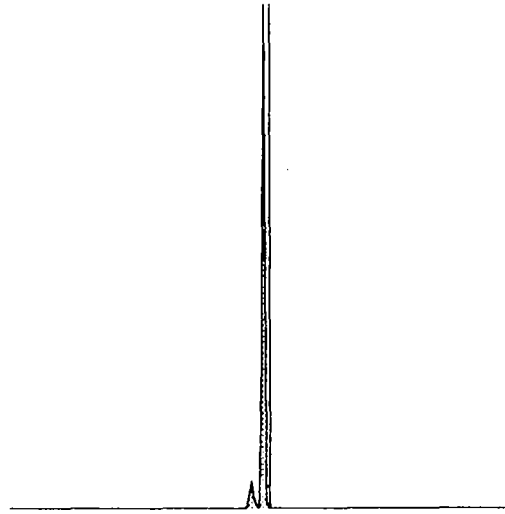


Figure 12: Sous ($HD1$), exemple 6.4: densité conditionnelle à l'instant $t = 6.60$

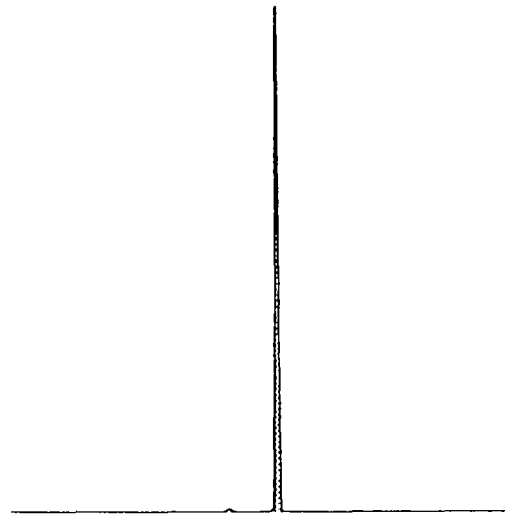


Figure 13: Sous ($HD1$), exemple 6.4: densité conditionnelle à l'instant $t = 6.86$

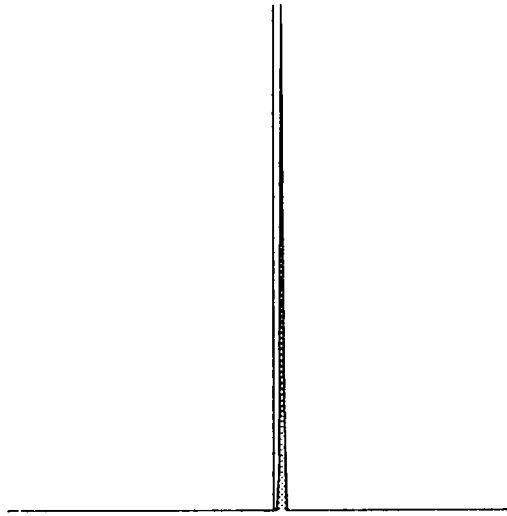


Figure 14: Sous $(HD1)$, exemple 6.4: densité conditionnelle à l'instant $t = 6.88$

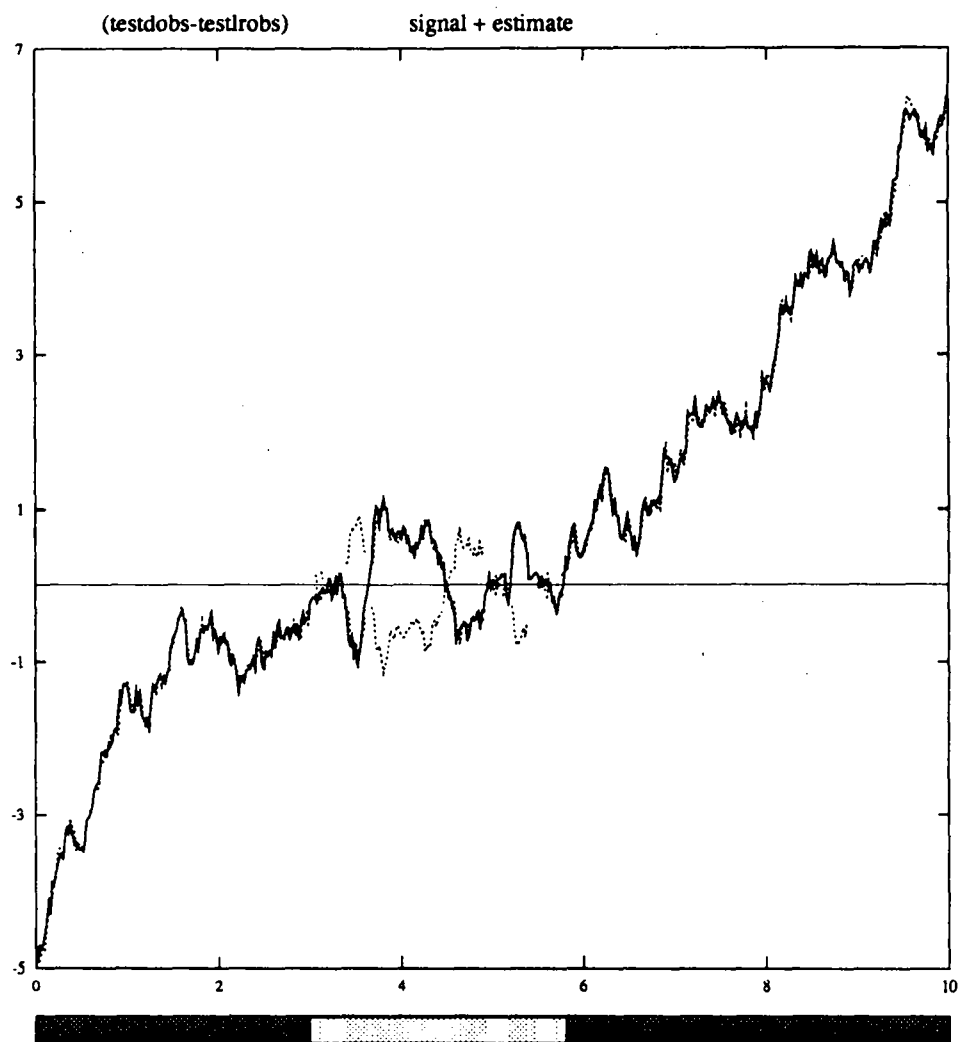


Figure 15: Exemple 6.6: test de détection des passages à zero basé sur les observations et test de décision du signe basé sur la variation quadratique.

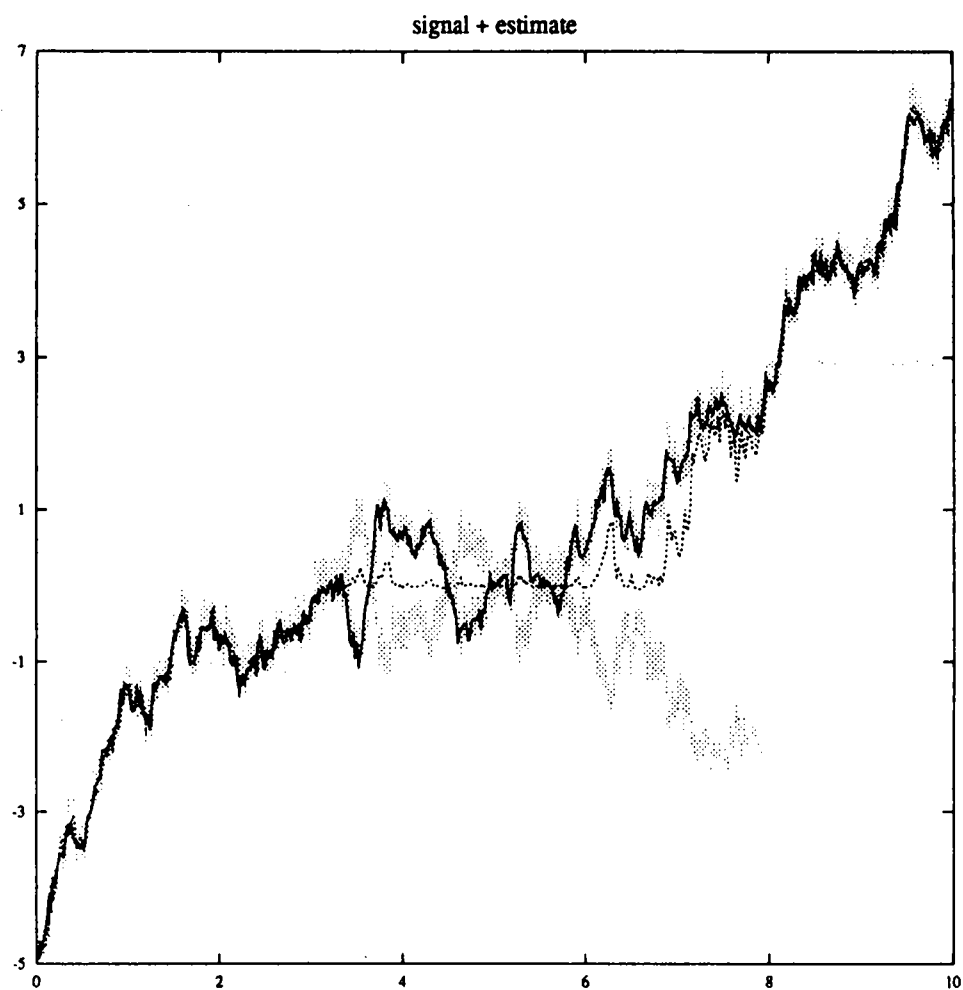


Figure 16: Exemple 6.6: solution de l'équation de Zakai (le niveau de gris correspond à une région de confiance à 95%).

ISSN 0249 - 6399